LOCALIZATION

USE CASE

CLEAR® VISION CLOUD

PFT's native media recognition AI platform CLEAR® Vision Cloud helps solve real-world business problems of TV networks, studios, and Streaming platforms because of its perfect combination of technology and consulting.

CLEAR® Vision Cloud is a fusion of homegrown AI engines (over 35 & counting) and best-of-breed AI models enveloped by PFT's unique Machine Wisdom layer. The Machine Wisdom makes a higher-level sense of details identified by the basic AI engines to increase accuracy and actionability. Think of it as an AI platform with a human-like brain (or close to it).

At PFT, our media services team has been leveraging AI to deliver customer projects for over two years now. This experience has helped us learn two things:

- AI delivers in unique M&E business contexts, and

- AI models can be tweaked to provide more accurate outcomes

As a result, CLEAR® Vision Cloud produces accurate data and makes it actionable.

Of the many successful use cases PFT has championed in the localization space, here are two that can add immense business value to content enterprises.

Use Case 1

Automatically generate, transcribe, trans-create subtitles in over 60 languages.

Localization is a laborious and error-prone process necessary to make your content available to audiences worldwide with different language preferences. Subtitle script creation involves transcription/translation, syncing it with video and audio, and re-timing. This manual work consumes time. In addition, at times, due to the poor quality of audio, the subtitle script may get some words or phrases wrong. Spelling and grammar errors while typing the captions are also common. Building an AI-enabled language transcription/translation tool that can deliver substantial business benefits is an arduous task. Even if the solution offers 95% accuracy, the captioning specialist has to review the entire process, making the automation redundant as time and cost benefits are nullified. Furthermore, with subtitling in 60+ languages, clients specifying parameters like characters per line, the number of lines and reading speed, and subtitle, the sync with video paramount. Nevertheless, there are quite a few things AI should get right most of the time to make AI work for studios, broadcasters, OTT players and, station groups.

And we have tried to achieve just that!

When PFT deployed the language transcription/translation tool powered by AI internally, it did not initially deliver even 50% accuracy. Our teams pushed back because the AI engine took longer TAT (turn-around time) and made more errors than during manual execution. This meant our operators had to spend extra time and effort to QC the AI delivered outcomes leading to natural frustration.

Through integrations with AWS, Google, and Microsoft and machine wisdom built in-house, including Indian regional languages, we substantially improved the accuracy of transcripts over a period of time. In addition, the content processed by CLEAR® Vision Cloud gets preloaded for QC, drawing attention to a text whose confidence is low or where captions are missing. This ensures that even at a 70-85% accuracy, substantial time and cost savings are achieved as the caption specialist does not have to review the entire process manually.

Apart from auto-generating the captions ensuring a very high level of accuracy, CLEAR® Vision Cloud offers side-by-side comparison aiding the caption specialist to fill in the gaps quickly.

- English language accuracy – 80-90%; other languages – 70-85%

- AI will accurately match the video/audio to the subtitle and provide a side-by-side result.

- The result – the user can now see the auto-generated transcribed subtitles against each shot. As the characters speak on the left pane of the video, the text gets highlighted on the right pane of the video.

- Wherever the system cannot auto-generate the subtitles, it gives the user a blank (length of the blank changes depending upon the length of the word), so the user can listen in to that particular piece of audio & manually type in the content.

- There is also a trans-creation option if required.

Advantages

CLEAR® Vision Cloud has ensured far greater accuracy apart from reducing time and effort involved in Localization by tweaking the AI models to suit specific M&E requirements.

- Automatic subtitle generation plus time-coded subtitles against each shot means the user will no longer have to type the subtitles manually.

- The subtitles thus generated are automatically synced with the video.

- The user gets a side-by-side view of content and along with the transcript. So the user no longer will have to toggle between video and script.

- Dictionary (road-map feature) also helps users spell-check and context check to correct subtitles detected or keyed in, making it easy for the user.

- The SRT file thus generated can be exported to CLEAR® for downstream activities.

Use Case 2

Automatic subtitle re-timing with 100% accuracy.

Global syndication through digital distribution is an essential monetization strategy for content enterprises. However, when the content is distributed in different markets across various channels, they often need to undergo:

- specific edits to fit slot duration

- or provide cuts and zooms to remove certain compliance

- or content moderation issues and add some extra frames with mandatory frames, disclaimers, among others.

These edits result in subtitles going off sync with the actual dialogues and warrant a re-time coding of subtitles. Unfortunately, subtitle re-timing is a manual process and takes a lot of time to ensure accuracy. As a result, it severely slows down the syndication workflows.

An AI-enabled comparator tool that can compare video frames from source and edited versions accurately is challenging. If the comparator fails to detect time codes of cuts and inserts accurately, it leads to erroneous re-timing of the subtitles.

When PFT deployed the AI-led comparator tool for subtitle re-timing for clients, it did not initially deliver even 50% frame accuracy. This meant a QC operator had to review the entire footage frame by frame for errors and fix them on the CLEAR® subtitling tool. In addition, the low level of frame accuracy invariably meant longer cycle time. But as the machines started learning from manual QCs, their ability to compare pre and post edit masters and identify cuts, edits, and inserts frame accurately dramatically improved, leading to much faster automatic subtitle re-timing.

- AI-led Video Comparator compares both source and edited versions of the videos and identifies matched and un-matched segments.

- Frame-accurately provides time codes of cuts and inserts in an edited video.

- Automatically re-times the subtitles on the edited video.

Advantages

The AI-led comparator can dramatically enhance frame accuracy to near 100% and considerably reduce subtitle re-timing cycle time, leading to better efficiencies and economics.

- Comparing the source with edited versions along with automatic re-timing of the subtitles means all the user will now have to do is a quick round of QC to make sure the subtitles are perfectly synced in the edited version of the video.

- No need to traverse frame-by-frame to check its accuracy; errors, if any, will automatically be flagged.

To learn more, visit **www.primefocustechnologies.com/vision_cloud**