Media Processing Service

Application Scenario and Practical Tutorial

Product Documentation

Contents

Application Scenario and Practical Tutorial
Short Drama Translation Scenario

Last updated：2025-03-21 15:25:03

Overview

Tencent Cloud's media processing products provide solutions for short dramas' global expansion, aiming to help creators and producers introduce them to international markets. This document introduces the functionalities and usage methods to achieve internationalization and localization of short dramas quickly.

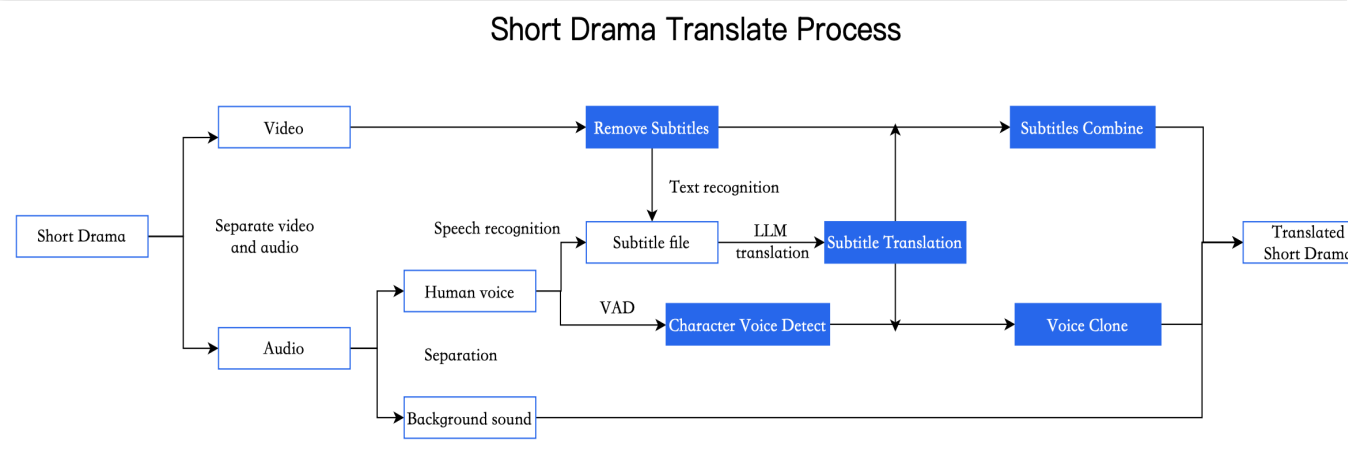
Left: Original video with CN text and Chinese voice; Right: Generated bi-lingual subtitles and speech after processing; Note: Image generated by Tencent Yuanbao AI

Requirements for Short Drama Clients

The challenges faced in the promotion of a vast array of videos include inefficient manual processing, the necessity to automatically recognize and generate multilingual subtitles, the erasure or replacement of items within the video, as well as substituting the original language dialogue with the target market language and re-dubbing to enhance the viewing experience.

Our Solutions

Short dramas seeking international exposure urgently require solutions for subtitle translation, dubbing conversion, and multilingual translation. We offer a comprehensive one-stop solution for the translation and production of short dramas.



Feature	Description
One-Stop Short Drama Translation and Production (Subtitle Removal + Voiceover & Translation + Audio Synthesis & Replacement)	This feature offers a comprehensive solution for original subtitle erasure, subtitle & voice translation, and multilingual subtitle & voice replacement. Utilizing this capability, the input video will be processed in the following sequential steps: 1. Identify the subtitles within the original video's visuals, followed by their erasure. 2. Recognize the original video's audio to generate subtitle files, which are then translated. The translated subtitles are subsequently: 2.1. Rendered onto the original video's subtitle positions. 2.2. Used to replace the original video's soundtrack through voice synthesis. 3. Synthesize a new video. **Example**

Left chart: Original Chinese short video; Right chart: The processed output video has English subtitles and English dubbing.

Distinct capabilities can also be integrated independently, as detailed below:

Feature	Description
Erasure of watermarks, logos, subtitles	Remove watermarks, logos, subtitles, and other elements from videos, including blurring and seamless removal options. Supports erasing many common watermarks and offers custom training for others at an additional cost.
Facial and License Plate Sensitive Information Blurring/Mosaic	Apply blurring or mosaic effects to sensitive information such as faces and license plates within video frames.
Human voice and background sound separation	Separation of human voice and background sound.
Text recognition	Recognizes text in a video, including vertically oriented text, and automatically extracts keywords from the text.
Speech recognition	Quickly recognizes the speech in a video and converts it to text based on deep learning. You can specify custom keywords and locate the time points in the video at which the keywords are spoken.
Subtitle Translation	Translate original subtitles to generate multilingual subtitles, catering to the needs of audiences from diverse linguistic backgrounds. Supports the configuration of a terminology database to enhance translation accuracy.

Incorporating Subtitles into Videos	Automatically generated subtitle files are embedded into videos through transcoding, allowing for customization of style, font, size, color, and background color.。
Text-to-Speech Conversion	Extract subtitle files to create new multilingual audio files.
Replace the original video's audio track	Replace the original video audio with generated multilingual audio to enhance the viewing experience for local audiences.
Digital Rights Management (DRM)	For copyrighted content, we offer support for commercial DRM and HLS private encryption. Currently, we support DRM encryption formats including FairPlay and Widevine.

Image Quality Improvement Scenario

Last updated：2025-03-21 15:24:37

Overview

Tencent Cloud Media Processing enhances video quality in on-demand and live streaming, boosting user retention and commercial conversion. It offers technical solutions across gaming, talent shows, education, and film, adapting to different content needs with its technical expertise and industry experience.

Key Requirements

Enhance video quality by rectifying visual imperfections, thereby presenting a superior viewing experience.
Capable of intelligently processing on demand, reducing costs and increasing efficiency.

Our Solutions

Abundant Audio and Video Enhancement Capabilities

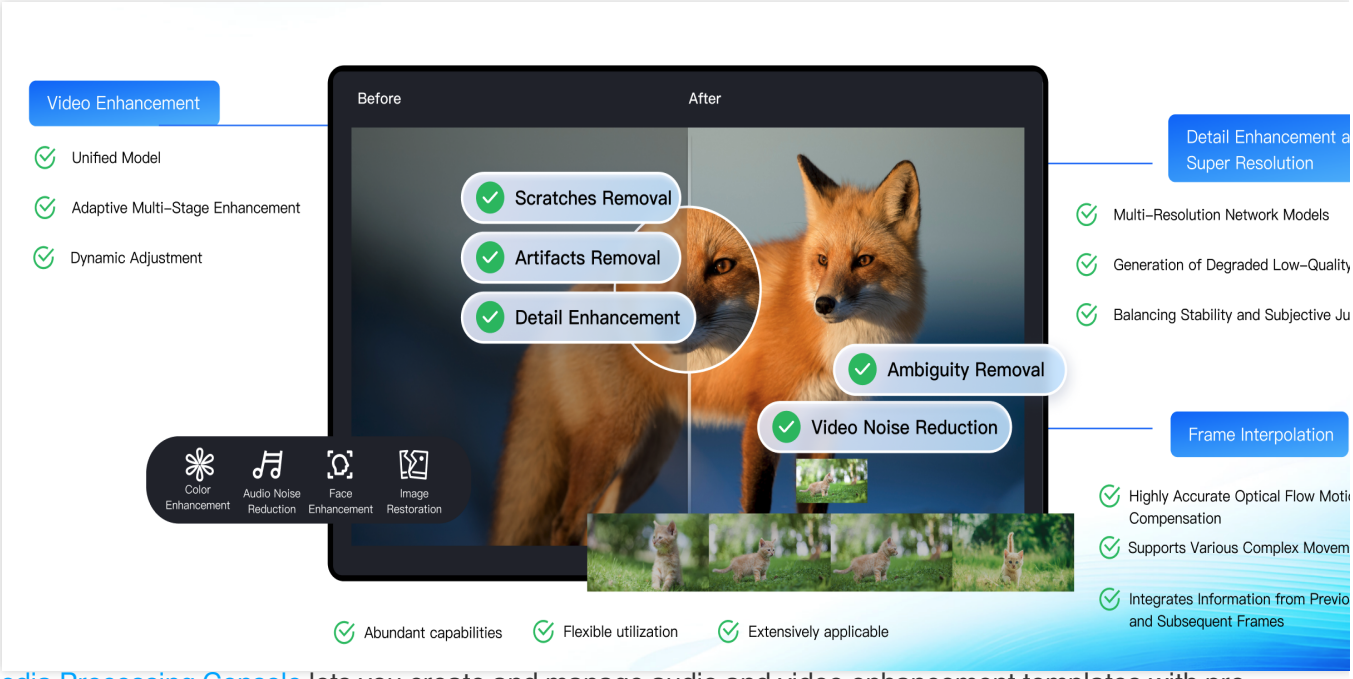
Tencent Cloud Media Processing provides various audio and video enhancement options. Users can use single or multiple features for better results.

Enhancement Type	Features	Description
Video Enhancement	Super Resolution	Super-resolution can identify the content and contours of the video, reconstruct the details and local features of the video in high definition, converting low-resolution videos into high-resolution ones, suitable for scenarios like old film restoration.
	Low-light Enhancement	Due to environmental conditions and limitations of the camera hardware, some scenes may suffer from a lack of brightness and contrast, resulting in dark images or missing details. Activating low light enhancement significantly improves the details and contrast in dark areas, enhancing the subjective quality of the human eye.
	HDR	Supports HDR10, and HLG, offering a wider color gamut and more color details, providing higher-quality video content.
	Comprehensive Enhancement	Through AI's comprehensive analytical capabilities, it automatically balances the texture content in the image, enhancing key details while

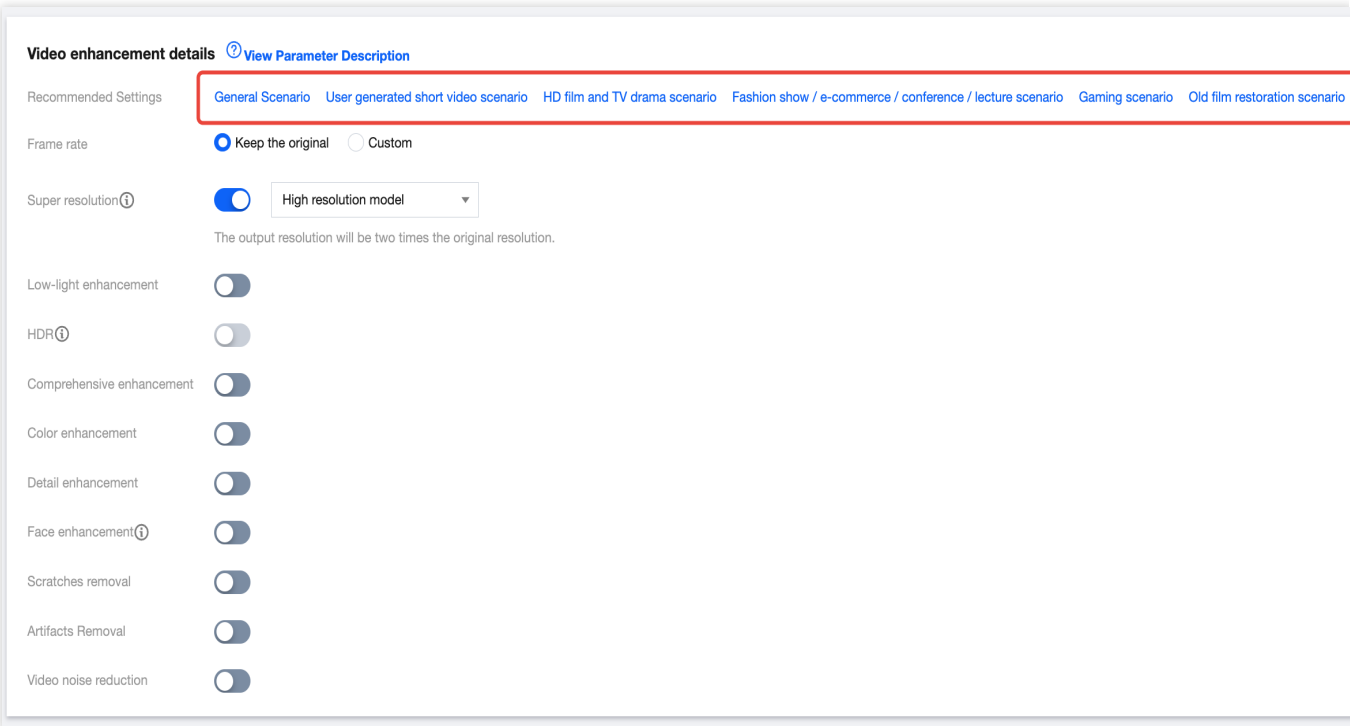
		removing compression artifacts and jagged edges, thereby improving the overall subjective perception of the image.
	Color Enhancement	Color enhancement makes the image closer to real colors and enhances them to some extent to meet the preferences of the human eye.
	Detail Enhancement	Detail enhancement focuses on the details in the video that require attention (e.g., the grass on a sports field), making the image content clearer and richer.
	Face Enhancement	Enhance the areas of the video that the human visual system particularly focuses on, such as faces, making the details in these areas clearer and improving subjective perception.
	Scratches Removal	Scratch removal can repair scratches and snowflake spots in the video, restoring damaged content.
	Artifacts Removal	Due to multiple compressions of the video during transcoding or multiple transcoding processes, block effects, ringing effects, chroma bleeding, and mosquito noise are introduced, causing distortions that affect the visual effect. De-compression distortion effectively repairs distortions introduced by encoding.
	Video Noise Reduction	Random noise may be introduced during film shooting due to the camera and environment. This service offers denoising, eliminating random noise in the image without losing detail.
Audio Enhancement	Audio Noise Reduction	Removes device noise, environmental noise, etc., suitable for scenarios such as recording classes and post-production of outdoor shooting.
	Audio Separation	Separates human voices from background sounds, or singing voices from accompaniment in audio-video files, creating independent audio materials for post-production artistic processing.
	Volume Equalization	Loudness Normalization: Maintains a consistent overall loudness level, making the playback sound similar in volume, avoiding issues of being too loud or too quiet, and providing a better auditory experience. Volume Leveling: Smoothens overly loud audio segments, avoiding sudden volume changes, and providing a more stable auditory experience.
	Audio Improvement	Noise removal: Reduces unwanted noise or interference in the audio, improving the quality and clarity of the audio. De-essing (Sibilance Suppression): Sibilance refers to the sharp, piercing sounds in audio, often produced when the sound source is

		close to the microphone. Suppressing sibilance aims to reduce or eliminate this unnatural sound, thereby improving audio quality.

Refer to the figure below:



The Media Processing Console lets you create and manage audio and video enhancement templates with pre-configured and customizable parameters for various scenarios, helping you quickly meet your business needs.



On-Demand Processing Solutions

Tencent Cloud's audio and video services offer enhancements and on-demand "Quality Inspection + Transcoding & Enhancement". It allows for cost optimization by inspecting video quality before enhancement. Based on inspection results, a suitable enhancement template is chosen. The improvement is measured by comparing video scores before and after processing.

Feature	Description	How to use
Quality Inspection	Quality Inspection includes: Format quality inspection: Detects format issues such as DTS, PTS problems, resolution changes, sampling rate changes, frame loss, and duplicate frames. Supports detecting the image quality of videos, with the specific inspection items as follows: JitterResults, BlurResults, AbnormalLightingResults, CrashScreenResults, BlackWhiteEdgeResults, NoiseResults, MosaicResults etc. No-Reference quality assessment	Refer to the Media Quality Inspection Integration documentation.
Audio and Video Enhancement	In on-demand processing scenarios, the focus is on fixing audio and video issues identified during quality checks through targeted interventions. Low-light Enhancement Comprehensive Enhancement Face Enhancement Audio Noise Reduction Videos without issues don't need enhancement, which could help to reduce costs.	Refer to the ProcessMedia documentation.

Audio/Video Cost Optimization Scenario

Last updated：2025-03-21 15:24:03

Overview

Live commerce, online education, remote conferencing, short videos, short dramas, and AI creation trends lead to more video content. Providers must balance efficiency and cost. Cost optimization is key for sustainability and competitiveness. Tencent Cloud Media Processing helps reduce processing and storage costs.

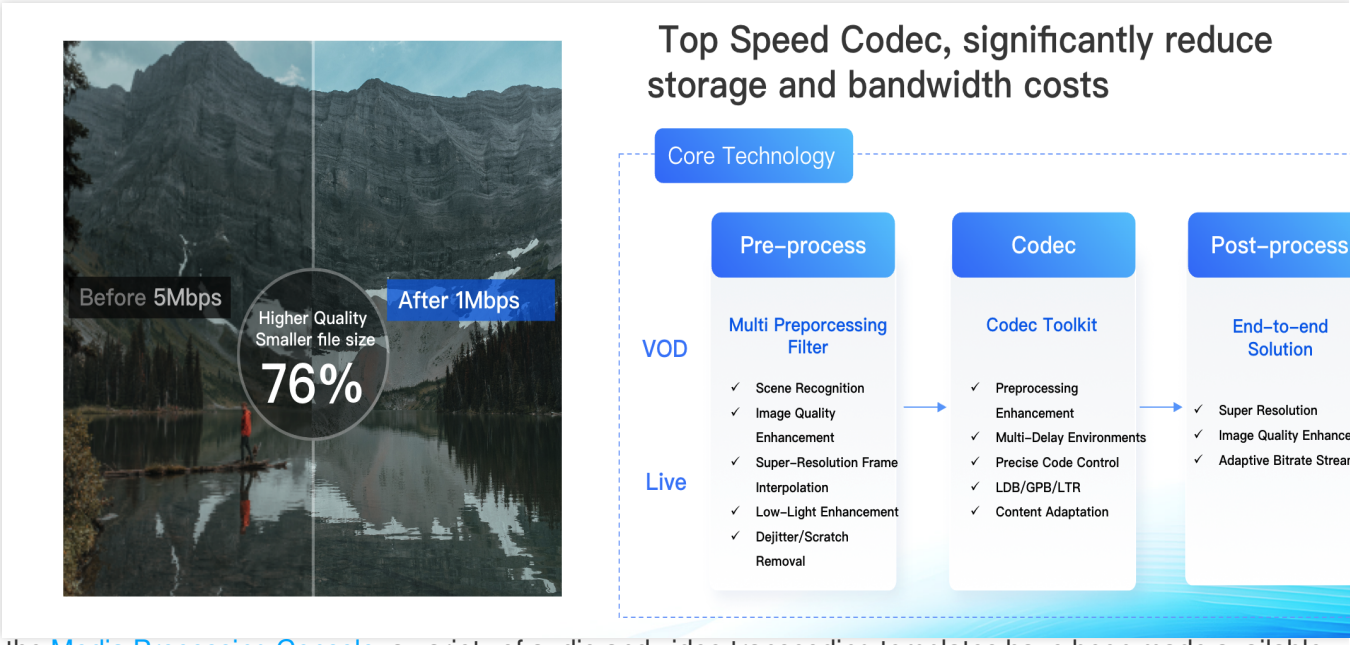
Requirements

The surge in video content has raised storage and data transmission costs, requiring tech solutions to reduce expenses.
Balancing the number of activated features is crucial in audio and video processing to avoid minimal optimization or high costs.
Differentiating video processing based on specific issues could maximize cost benefits.

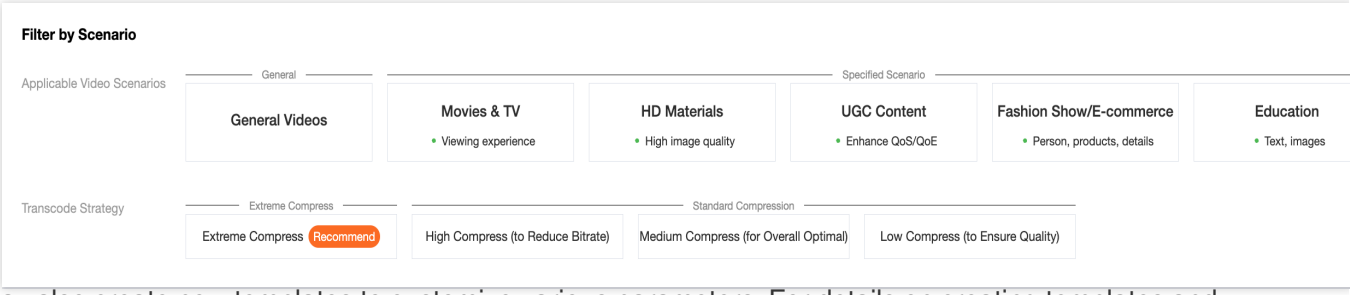
Our Solutions

Leading Encoding Compression Capabilities

Supports encoding standards such as VP8, VP9, H.264, H.265, AV1, AVS3, H.266, etc., and supports real-time encoding for 4K and 8K, providing an ultra-high-definition and smooth video experience. While ensuring the subjective quality of the video, it saves 50% of bandwidth costs. For the past three years, it has consistently won first place in the MSU Video Coding Competition; it has also achieved the best overall performance in the Streaming Learning Center's encoding evaluation. You can experience the transcoding effects on our online demo page.



Within the Media Processing Console, a variety of audio and video transcoding templates have been made available for your convenience. Each template comes pre-configured with a set of transcoding parameters, allowing you to swiftly select the appropriate transcoding template tailored to your specific business scenario.



You may also create new templates to customize various parameters. For details on creating templates and configuring parameters, please refer to Audio and Video Transcoding Templates. For information on the costs associated with audio and video transcoding, please visit Billing Instructions.

On-Demand Processing Solution

Tencent Cloud's audio and video services provide high-quality transcoding and an on-demand "Quality Inspection + Transcoding & Enhancement" solution. To optimize costs, videos can be inspected before transcoding, allowing for a suitable template selection based on the inspection results.

Feature	Description	How to use
Quality Inspection	Media Quality Inspection Supports: Format quality inspection: Detects format issues such as DTS, PTS problems, resolution changes, sampling rate changes, frame loss, and duplicate frames.	Refer to the Media Quality Inspection Integration documentation.

	Video & Audio Content Quality Inspection: JitterResults, BlurResults, AbnormalLightingResults, CrashScreenResults, BlackWhiteEdgeResults, NoiseResults, MosaicResults etc. No-Reference quality assessment	
General transcoding/TSC transcoding/Adaptive bitrate streaming	1In on-demand processing scenarios, media assets are first inspected for quality. Based on this inspection, videos with format problems or high bitrates are transcoded, while those without issues skip this step, reducing transcoding costs.	Refer to the ProcessMedia documentation.

Evaluating after Transcode

Upon completion of the transcoding process, users can also quantitatively assess the transcoding results through the transcoding evaluation feature. Tencent Cloud Media Processing offers comprehensive transcoding evaluation capabilities, supporting VMAF, PSNR, SSIM, and VMAF-NEG scoring for videos from various sources and in multiple formats.

Classification	Feature	Description	How to use
On-demand video	Video quality evaluation	Adds the original video and the comparison video to perform video quality evaluation. Supports evaluation methods including VMAF, PSNR, SSIM and VMAF-NEG. Supports customizing the selection of a time period or range of frames for evaluation.	Refer to the User Guide documentation.
	BD-Rate comparison evaluation	Selects a Media Processing Service template and evaluates the differences in video transcoding quality of different templates at various bitrates. Supports evaluation methods including VMAF, PSNR, SSIM, and VMAF-NEG. Supports customizing the selection of a time period or range of frames for evaluation. Supports comparing evaluation scores at specified bitrates or comparing bitrates at a specified CRF (video quality score).	Refer to the User Guide documentation.
Live stream	Image	Supports real-time comparison and monitoring of	Coming soon.

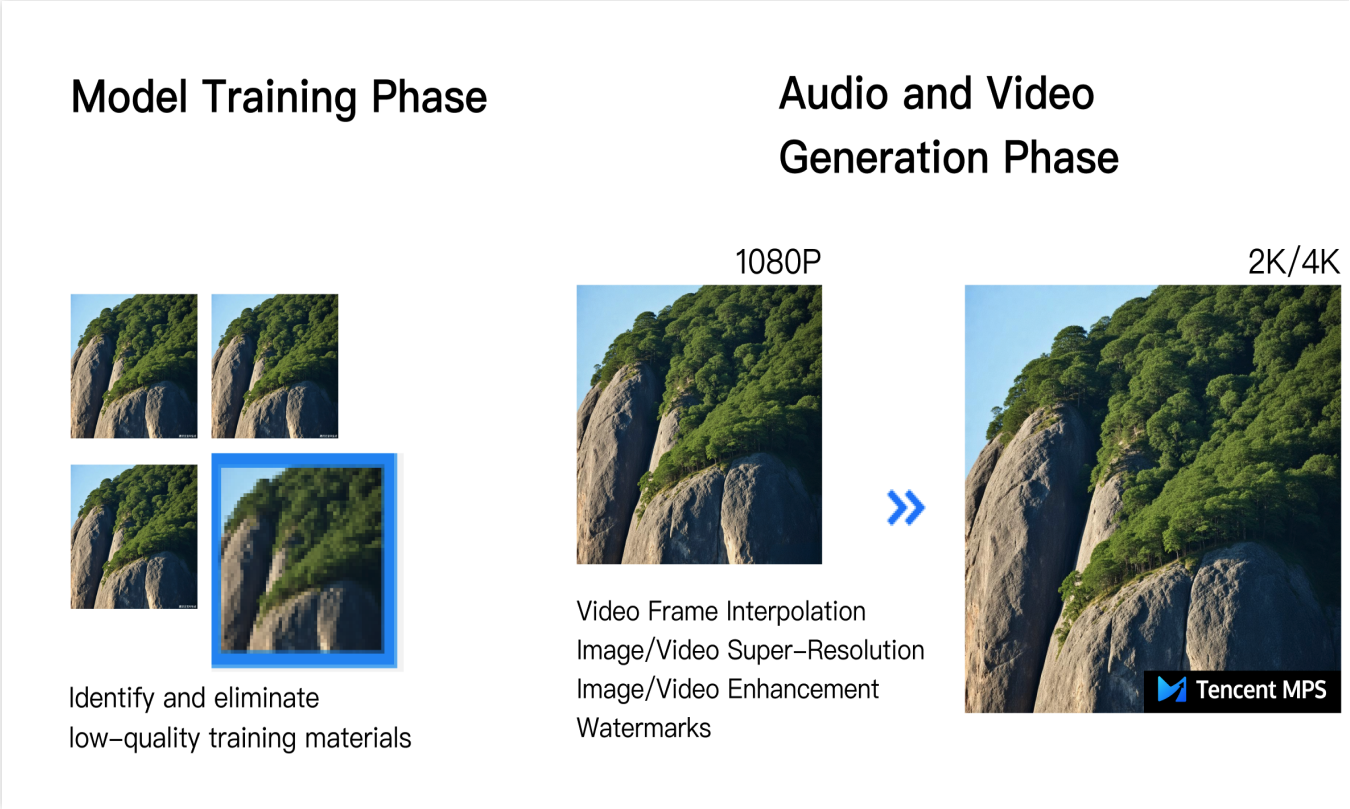
	quality	image quality and bitrate changes before and after live stream transcoding.	

AI-Generated Content Scenario

Last updated：2025-03-24 10:57:11

Overview

Tencent Cloud's services enhance AI-generated content by integrating quality inspection and advanced imaging technologies, leading to smoother and more lifelike outcomes.



Scenarios

Quality Discrimination during the Collection and Cleansing Phase of Training Materials

Requirement: Filter out low-quality videos.

Solution: By using Tencent MPS quality inspection, videos with format and visual problems can be removed to improve training video quality.'"

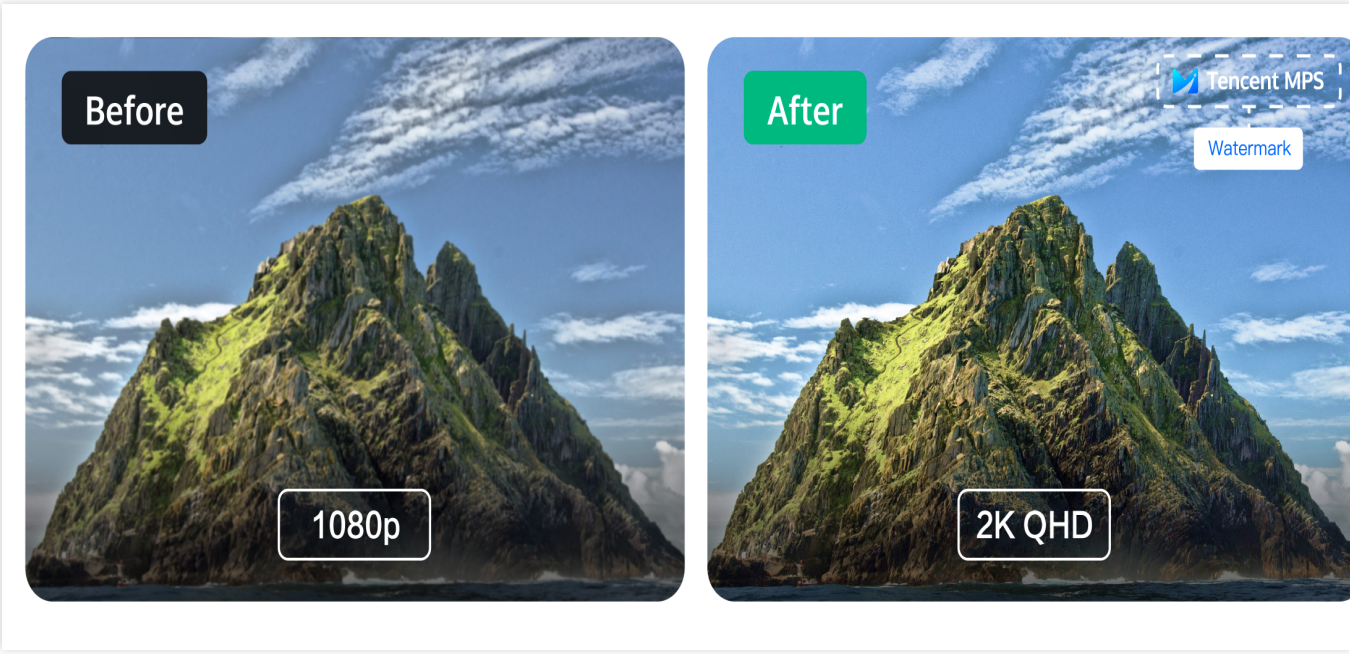
Enhancing Video Resolution, Frame Rate, and Color Fidelity

Requirement: There is a need to enhance the resolution, frame rate, and color effects of the generated video to achieve a clearer and smoother visual experience.

Solution: In the context of video generation, the video super-resolution capability can further improve the clarity of AI-generated videos and achieve a smoother visual effect through intelligent frame interpolation.

Meeting Compliance and Information Traceability Requirements

Requirement: Incorporate essential watermarking into the generated results to ensure traceability and compliance with regulatory requirements.
Solution: Implement the addition of emergent and concealed watermarks to the outputs of large-scale models, ensuring traceability and compliance of information.



Feature Overview and Integration Method

Feature	Description	How to use
Quality Inspection	Media Quality Inspection Supports: Format Quality Inspection Video & Audio Content Quality Inspection No-Reference quality assessment During the material cleansing phase of model training, this feature is capable of detecting issues related to video formats, as well as identifying problems within the visuals such as color distortion, low lighting, vignetting, abnormal contrast, blur, and mosaic effects. By conducting quality inspections and eliminating problematic videos, the overall content quality of the training materials will be significantly enhanced.	Refer to the Media Quality Inspection Integration documentation.
Video	The feature improves video by deblurring, reducing noise,	Refer to the

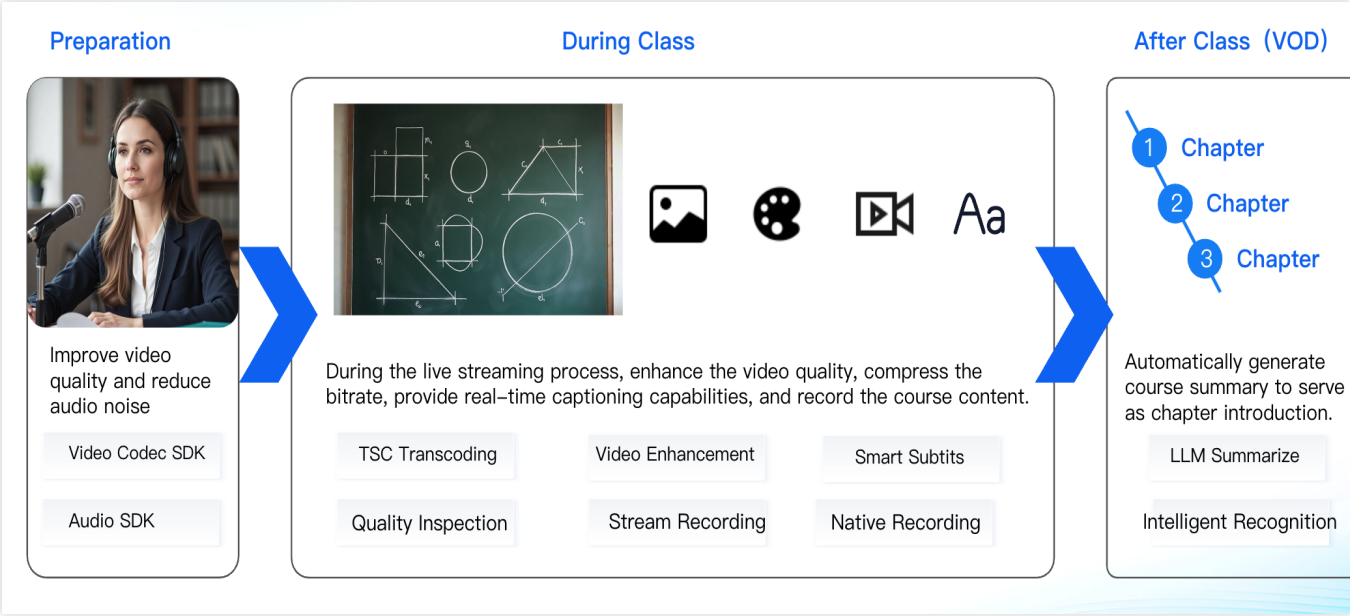
Enhancement	enhancing details, text, color, faces, and resolution, and interpolating frames. It can upscale 720P or 1080P videos to 2K or 4K, further improving quality through frame interpolation, detail, and color enhancement..	ProcessMedia documentation.
Image/Video Watermark	Supports the addition of both visible watermarks and invisible watermarks to videos, fulfilling requirements for information traceability and compliance. Supports customization of watermark styles, including watermark images, positions, and sizes.	

Online Education Scenarios

Last updated：2025-03-21 15:23:21

Overview

The online education industry is transforming due to technology. With a growing user base and new tech, the focus is on maintaining teaching quality with limited bandwidth, cutting costs, and integrating new technologies. This document covers Tencent Cloud's Media Processing for education, helping users apply tech solutions to save costs and boost efficiency.



Scenarios

Clear and Stable Live Streaming Performance

Requirement Description: A clearer visual and stable audio quality are crucial for students to see blackboard content and understand courseware information without hindrance.
Solution: By leveraging terminal video encoding and audio noise reduction capabilities, we enhance the audio and video quality for teachers in online classes, ensuring clear visuals and stable sound. Utilizing ultra-fast transcoding and fast live streaming capabilities, we deliver low-latency, high-definition live streaming experiences.

AI Enhances Educational Quality

Requirement Description: AI should be used in and after class to improve education quality.

Solution: Employ the intelligent subtitle feature to display the instructor's lecture content as real-time captions, with the large model summary function, to generate a course summary and key points overview at the conclusion of the session, thereby enhancing the teaching outcome.

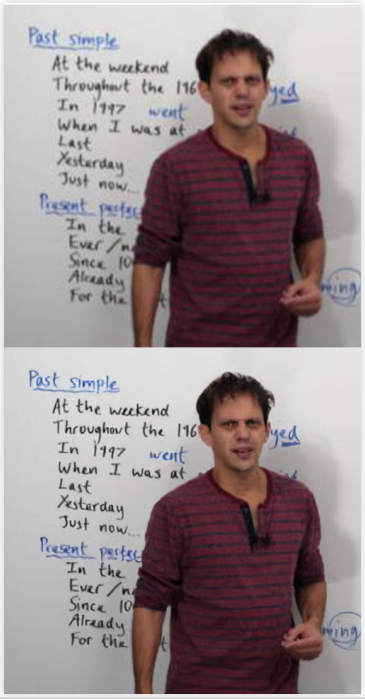
Enhancing the Broadcasting Experience for Instructors

By enhancing the encoding performance on the teacher's end and improving the audio noise reduction capabilities, the quality of audio and video streaming from the teacher's broadcasting interface can be optimized.

Feature	Description	How to use
Terminal Video Codec SDK	Tencent's Top Speed Codec (TSC) terminal video encoder is designed for scenarios requiring low computational power, low latency, and high-quality image on the terminal side. Compared with hardware encoding, its advantages include: 1. Stable, reliable, and quick to start. 2. At the same quality level, it saves bitrate, enhances transmission stability, reduces downlink distribution bandwidth, and saves on storage costs. 3. At the same bitrate, it improves image quality and enhances user experience. 4. A rich set of features meets diverse business needs, such as using Regions of Interest (ROI) encoding to improve the image quality in the face region and dynamically adjusting encoding configuration to adapt to network fluctuations.	Step One: It is requisite to activate a trial License via the console Activate Test License or Procure an Official SDK License; Step Two: For technical integration, please refer to the Ultra HD Terminal Encoding SDK Integration process.
Terminal Audio SDK	The Terminal Audio SDK, integrated on the client side, significantly enhances audio quality, eliminating echo and noise. The technology employed includes: 1. Adaptive Noise Suppression, Eliminate background noise components in the audio signal and reduce interference to enhance the perceptual quality of the voice. 2. Acoustic Echo Cancellation mainly addresses the echo problem in audio communication. 3. Automatic Gain Control responsible for adjusting the volume during the transmission of audio signals. 4. The terminal audio quality inspection capability enables the scoring of audio input quality, promptly identifying audio issues and providing alerts.	Step One: It is requisite to activate a trial License via the console Activate Test License or Procure an Official SDK License; Step Two: For technical integration, please refer to the Ultra HD Terminal Encoding SDK Integration process.

Enhancing the Viewing Experience for Students

Feature	Description	How to use
Top Speed Codec(TSC) Transcoding	TSC transcoding is an "Upgrade" version of standard video transcoding, capable of adaptively optimizing various videos to deliver a higher definition viewing experience with lower bandwidth. While ensuring or enhancing image quality, it saves more than 50% in bandwidth costs.。	Refer to the ProcessMedia documentation.
Quality Inspection	Media Quality Inspection Supports: Format Quality Inspection Video & Audio Content Quality Inspection No-Reference quality assessment This feature detects live stream issues, alerting operators for quick anomaly identification. For video-on-demand scenarios, it detects and scores content, integrating transcoding and enhancement features for problem resolution.	Refer to the Media Quality Inspection Integration documentation.
Video Enhancement	Leveraging the industry-leading AI processing models within MPS, alongside a plethora of applications in various business scenarios, the enhancement feature significantly elevates audio and video quality. Supports distributed real-time image quality enhancement, artifacts removal, noise reduction, color enhancement, detail enhancement, and face enhancement. In educational settings, it is possible to enhance video imagery by optimizing color representation, refining facial details, and improving the clarity of written content on boards.	Refer to the ProcessMedia documentation.

		
Cloud-based Stream Record	Traditional recording only captures visual effects, not audio and video content on the client side. To record exactly what is seen, it's necessary to use cloud-based Windows or Android to stream via Tencent Cloud Live Broadcasting, integrating with the recording module to create files. This feature can be utilized to generate course replays and teacher's presentations.	Please contact the business team to have Beta test.
Smart Captions and Subtitles	The Smart Captions and Subtitles feature recognizes voices from videos or live streams in real-time, converts speech to subtitles, and translates them into various languages like Chinese, English, and Japanese-Korean. It's useful for live subtitles, video transcription for global distribution, and extracting subtitles from old films for restoration. The feature supports configuring HotWord for more accurate translations. In educational settings, the subtitle feature can be utilized to generate subtitles in real-time during live classroom sessions conducted by teachers; it can also be employed to add subtitles to on-demand videos, thereby enhancing the viewing experience for users.	Refer to the Subtitle Generation and Translation documentation.
LLM Summarize Tutorial	Leveraging the foundation of intelligent recognition, the large model summary is capable of generating concise summaries for video content. It automates the creation of key points and comprehensive summaries of course content.	Refer to mps.live for free trial. Refer to the LLM Summarize Tutorial documentation.