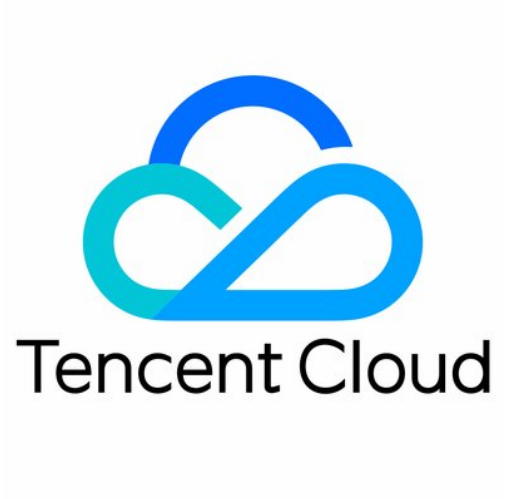
Media Processing Service

Product Introduction

Product Documentation



Contents

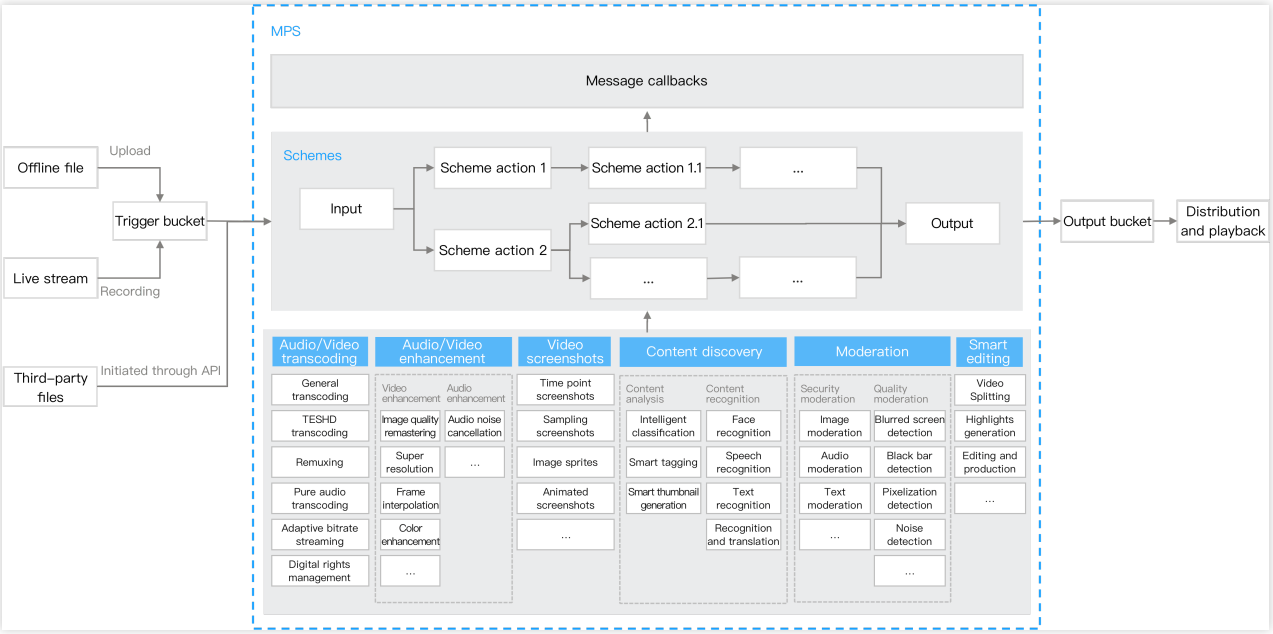
Product Introduction
Overview

Last updated：2022-08-10 15:36:52

MPS is a cloud-based audio/video processing service. Built on Tencent's many years of experience in audio/video technologies, it provides ultimate encoding capabilities that enable you to play back audio/video files on various platforms at greatly reduced storage and bandwidth costs. It provides a rich set of features including video screencapturing, audio/video enhancement, content discovery, and content moderation, satisfying your video processing needs in a range of different scenarios.

Product Architecture

You can upload source video files to a COS bucket through the console, SDKs, or APIs. You can use the **scheme** mechanism of MPS to trigger the automatic execution of video processing tasks and send event notifications to CMQ. This helps you stay informed of the status of transcoding tasks. The architecture of MPS is shown below:



Features

Last updated：2022-08-10 15:36:52

MPS transcodes audio/video files to different bitrates and resolutions for smooth playback on various devices with different bandwidth options. It has the following features:

Audio/Video Transcoding

Transcoding is an offline task that changes the codec, resolution, bitrate, and other characteristics of an audio/video stream to suit different playback devices and network conditions. The benefits of transcoding include:

Feature	Description
Increased compatibility	A source video can be transcoded to formats (such as MP4) that are compatible with more types of devices for smooth playback.
Increased bandwidth adaptability	A source video can be transcoded for output in multiple definitions such as LD, SD, HD, and UHD. End users can select the most appropriate bitrate depending on their network conditions.
Improved playback efficiency	The moov atom can be moved from the end of an MP4 file to the beginning of the file, allowing the video to be played before it is entirely downloaded.
Reduced bandwidth consumption	With a more advanced codec (such as H.265), the bitrate of a video can be substantially reduced while retaining the original quality, which helps reduce the bandwidth consumption.

The parameters you can specify for transcoding include codec, resolution, bitrate, etc. For details, see the table below.

Category	Parameter	Description
Input	Container format	3GP, AVI, FLV, MP4, M3U8, MPG, ASF, WMV, MKV, MOV, TS, WebM, MXF
	Video codec	AV1, AVS2, H.264/AVC, H.263, H.263+, H.265, MPEG-1, MPEG-2, MPEG-4, MJPEG, VP8, VP9, RealVideo, Windows Media Video, QuickTime
	Audio codec	AAC, ADPCM, AMR, DSD, MP1, MP2, MP3, PCM, RealAudio, Windows Media Audio, Vorbis, AC-3
Output	Container format	Video: FLV, MP4, HLS (M3U8 + TS), MXF
		Audio: MP3, MP4, Ogg, FLAC, M4A

		Image: GIF, WebP
	Video codec	H.264/AVC, H.265/HEVC, AV1
	Audio codec	MP3, AAC, FLAC, MP2, Vorbis
Packaging	Delete video streams	If this is enabled, the transcoding result will contain only audio streams.
	Delete audio streams	If this is enabled, the transcoding result will contain only video streams.

Audio/Video Enhancement

By combining the image quality remastering and video enhancement modules with AI algorithms, MPS supports image noise removal, contour restoration, super resolution reconstruction, and other features while improving the resolution, making it suitable for various business scenarios such as UGC/PGC video quality improvement, digital remastering, and 4K video production.

Capability	Description
Image noise removal	Removes the random noise introduced from the camera and the environment during video recording while maintaining details of the video image.
Artifact (glitch) removal	Effectively repairs distortions caused by repeated compressions of videos during transcoding that compromise the visual quality, such as blocking artifacts, ringing artifacts, color contamination, and mosquito noise.
Banding removal	Repairs banding and snow caused by various factors that affect the film during video recording, storage, or transfer.
Detail enhancement	Makes the video image clearer by enhancing details which may have been compromised by the camera quality or during video saving or transcoding.
Overall enhancement	Uses AI-based analysis to improve the overall image quality in videos by balancing image textures, removing compression artifacts, and enhancing key details.
Super resolution	Enhances and restores details in low-resolution videos that can't meet today's requirements for a high definition. It uses an AI model to output high-resolution videos with clearer details.
Face enhancement	Uses face detection to enhance the detail and quality of faces in the video.
Color enhancement	Restores video color that may have been distorted due to camera problems or video storage and enhances the color to to make it more pleasing to viewers.

Low-light enhancement	Due to the environmental conditions and the hardware limitations of the camera, the video image of certain scenes may lack brightness and contrast, leading to loss of details in dark areas. This feature automatically recognizes scenes and adaptively enhances the video image to increase details and contrast in dark image areas and improve the image quality, especially in low-light scenes.
HDR	Converts general SDR videos to HDR videos. It can increase the color depth to 10 bits to get a wider gamut and display more color details, providing higher-quality video content.
Frame interpolation	Adds additional video frames between the original video frames to offer a smoother visual effect, improving image quality in older videos shot at a low frame rate and reducing lag and jitter.

Watermarking

Watermarking is an offline task that adds an image at the specified position of the video during video transcoding or screencapturing. MPS supports the following types of watermarks:
Static watermark: Non-animated watermark in PNG format, which can be the logo of a copyright owner or TV station, and is usually used as a copyright claim.
Animated watermark: Animated watermark in APNG format
MPS allows you to add multiple watermarks to a video or screenshot and specify the size and position of each watermark in the video or screenshot.
The parameters you can specify for watermarking include watermark type, aspect ratio, position, etc. For details, see the table below.

Parameter	Description
Type	The watermark type. Watermarks can be static or animated.
Position	The relative position of a watermark in the video.
ImageSize	The size of the watermark in the video.
ImageContent	Binary data of a watermark.

Video Screencapturing

Screencapturing is an offline task that captures a screenshot of a video at a certain point in time. MPS provides the following types of screenshots:

Time point screenshot: Screenshots taken at specified time points
Sampled screenshot: Screenshots taken at regular intervals
Image sprite: MPS can capture a set of screenshots of a video (subimages) at the specified time interval and splice them together to generate a large image (i.e., an image sprite).
The parameters you can specify for screenshot taking include screenshot format, aspect ratio, etc. For details, see the table below.

Time point screenshots

Parameter	Description
Format	The screenshot format (only JPG is supported currently)
Width	Screenshot width (px). Value range: 128-4096
Height	Screenshot height (px). Value range: 128-4096
FillType	The fill mode (`FillType`) specifies how the source video image processed when the aspect ratio does not match the specified aspect ratio of a screenshot. The following fill modes are supported: Scale to fill: Source video images are stretched to match the aspect ratio of screenshots. This may cause images to appear distorted. Black bars: The aspect ratio of source video images is retained, and the empty spaces are painted black. White bars: The aspect ratio of source video images is retained, and the empty spaces are painted white. Gaussian blur: The aspect ratio of source video images is retained, and Gaussian blur is applied to the empty spaces.

Sampled screenshots

Parameter	Description
Format	The screenshot format (only JPG is supported currently)
Width	Screenshot width (px). Value range: 128-4096
Height	Screenshot height (px). Value range: 128-4096
SampleType	How sampling intervals are measured. Sampling intervals can be measured in two ways: By percent: Intervals are measured by percent. For example, if `Interval` is set to 5 (%), 20 screenshots will be generated for a video. By time: Intervals are measured by time. For example, if `Interval` is set to 10 (sec), the number of screenshots generated will depend on the video length.

Interval	The sampling interval. If the interval measurement (`SampleType`) is by percent, this parameter is a percent value. If interval measurement is by time, this parameter is a time value (sec).
FillType	The fill mode (`FillType`) specifies how the source video image processed when the aspect ratio does not match the specified aspect ratio of a screenshot. The following fill modes are supported: Scale to fill: Source video images are stretched to match the aspect ratio of screenshots. This may cause images to appear distorted. Black bars: The aspect ratio of source video images is retained, and the empty spaces are painted black. White bars: The aspect ratio of source video images is retained, and the empty spaces are painted white. Gaussian blur: The aspect ratio of source video images is retained, and Gaussian blur is applied to the empty spaces.

Image sprites

Parameter	Description
Format	The format of the image sprite (only JPG is supported currently).
Width	The width of the subimage in an image sprite.
Height	The height of the subimage in an image sprite.
Rows	The number of image rows in a sprite.
Columns	The number of image columns in a sprite.
SampleType	How sampling intervals are measured. Currently, only sampling by time is supported.
Interval	The time interval for image sampling.

Note:

The result of multiplying `Width` x `Columns` (i.e., sprite width) should be within the range of 128-4096.

The result of multiplying `Height` x `Rows` (i.e., sprite height) should be in the range of 128-4096.

Animated screenshots

Animated screenshot generating is an offline task that converts a video segment to an animated screenshot such as in GIF or WebP format. An animated screenshot is a seamless cycle of continuous frames, which can deliver an animation effect with a small file size.

The parameters you can set for animated image generation include format, width, height, frame rate, etc. For details, see the table below.

Parameter	Description
Format	The format of the animated image (only GIF and WebP are supported currently).
Width	The animated image width. Value range: 128–4096 px.
Height	The animated screenshot height. Value range: 128–4096 px.
FPS	The frame rate. Value range: 1–60 fps.

Content Discovery

Content recognition

Based on the work of Tencent's research labs, content recognition recognizes various forms of video content such as people, speech, text, and frame tags and performs multidimensional structured analysis.

Recognition Type	Description
Face Recognition	Quickly recognizes facial information in a video based on deep learning and locates the frames in which a person is present as well as the position of the person’s face. You can use custom person libraries or call video AI-enabled public person libraries to recognize faces.
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.
Text recognition	Recognizes text in a video, including vertically oriented text, and automatically extracts keywords from the text.
Frame tag recognition	Uses deep learning to automatically recognize tags in the video frames captured at the custom frame capturing interval, and locates the tags in the video. Frame tags are divided into nine categories, such as people, landscape, artificial object, building, plant, animal, and food, covering various aspects of daily life. You can use custom tags based on the tag system. It has transfer learning capabilities, so you can customize classifiers simply by providing the raw user data. In this way, it meets the requirements of different types of users and makes the tag system more flexible.
Opening and ending credits recognition	Automatically recognizes and locates the time points of opening and ending credits of movies and TV series based on the video image

	characteristics, text, speech, and other information.

Content analysis

Analysis Type	Description
Category recognition	Recommends a category for the target video by analyzing the video content. Currently, it supports 19 categories, including food, travel, animation, and music. Custom categories are also supported as a paid feature.
Video tag recognition	Intelligently recognizes top five tags that best fit the video content based on Tencent's deep learning solution. It is suitable for video recommendation and search scenarios. You can customize the number of tags to be returned in the API.
Intelligent thumbnail	Automatically generates a file thumbnail based on characteristic information such as video image texture and scene recognition. It allows you to output static thumbnails quickly, making it easier to create thumbnails for videos and improving video click rates.

Smart Moderation

Smart moderation includes security moderation and quality moderation.
Security moderation uses AI to detect erotic, illegal, and non-compliant content in video images, audio, and text.
Quality moderation moderates the image frames and audio quality in live and on-demand videos. It supports 13 detection types such as blurred screen, black bar, pixelization, and noise. It also moderates and scores the overall quality of the video.

Moderation Type	Detection Type	Detection Item Description
Security moderation	Video image moderation	Moderates the video image to detect erotic and non-compliant content, specifically including: Erotic content detection `porn`: Pornographic content `vulgar`: Vulgar content `intimacy`: Content that displays intimacy `sexy`: Content that displays sexiness Illegal and non-compliant content detection `guns`: Weapons and guns `bloody`: Bloodiness `explosion`: Explosions and fires

		`violation_photo`: Banned icons
	Audio moderation	Moderates the speech in the audio based on the following: Erotic content detection: Analyzes speech in the audio to detect keywords related to erotic content. Illegal and non-compliant content detection: Analyzes speech in the audio to detect keywords related to illegal and non-compliant content.
	Text moderation	Moderates the text in video images, specifically including: Erotic content detection: Analyzes text in the video image to detect keywords related to erotic content. Illegal and non-compliant content detection: Analyzes text in the video image to detect keywords related to illegal and non-compliant content.
Quality moderation	Image quality	Detects the following in the video image: JitterResults: Jitter BlurResults: Blur AbnormalLightingResults: Low light or overexposure CrashScreenResults: Blurred screen BlackWhiteEdgeResults: Black bar, white bar, black screen, white screen, and solid color screen durations NoiseResults: Noise MosaicResults: Pixelization QRCodeResults: QR code
	Audio quality	Detects the following in the speech in the video: VoiceResults: Audio exceptions, including no sound, low volume level, and cracking

Smart Editing

Based on AI and audio/video technologies developed by Tencent, smart editing comprehensively discovers video content in various dimensions and supports smart highlights generation and video splitting to assist with video content production.

Capability	Description
Smart splitting	Performs structured analysis on the video content and intelligently splits the video into segments based on scene, speech, and text information. Currently, it is supported for news and ads.
Smart highlights generation	Based on video temporal/spatial characteristics matching, scene recognition, target detection, and other technologies, it automatically collects video highlights in various

	video scenes such as soccer, basketball, PlayerUnknown's Battlegrounds, and Honor of Kings. Custom video scenes are supported on a paid basis.
Editing and production	Allows you to clip and splice videos, convert images into videos, add roll images and text to videos, implement picture-in-picture, and edit audio.

Strengths

Last updated：2022-08-18 16:38:56

Process orchestration

MPS allows you to create custom schemes to orchestrate service processing for high numbers of media files and complete basic operations such as transcoding, screencapturing, and watermarking. You can use event callbacks to stay up to date with the progress of different tasks.

Top Speed Codec

MPS supports H.264, H.265, and AV1 encoding standards and ensures high video quality while reducing bandwidth costs by 50%. In addition, it supports real-time 4K and 8K encoding and delivers a smooth UHD viewing experience.

Accelerated transcoding

MPS is deployed in global regions and supports automatic scaling to flexibly handle a high number of concurrent transcoding requests. It supports distributed transcoding to deliver up to 30 times faster transcoding for long videos, allowing you to transcode and release videos very quickly.

Image quality remastering

By leveraging industry-leading video AI technologies, MPS supports digital remastering, SD to HD, and HD to 4K, allowing you to increase the quality and color richness of videos while greatly reducing video noise, glitches, and banding.

A Variety of Features

MPS supports various capabilities such as audio/video transcoding, audio/video enhancement, video screencapturing, content discovery, moderation, and smart editing. It provides a variety of template configuration capabilities, which enable you to customize the configuration as needed.

AI Technology

Based on Tencent Cloud's advanced AI technology, MPS provides diversified smart services for massive amounts of video content, including smart tagging, intelligent thumbnail generation, face recognition, speech/text recognition, smart editing, and security/quality moderation.

Use Cases

Last updated：2022-08-18 16:42:55

Video Websites

MPS covers all mainstream video formats and supports multiple resolutions and bitrates, enabling video websites to deliver a smooth and consistent user experience to users with different bandwidth options.

Strengths:

Provides an ultimate encoding experience and supports Top Speed Codec (TSC) transcoding, which greatly reduces the storage and bandwidth costs.

Supports adaptive bitrate streaming to enhance playback on various devices under different network conditions.

Offers a wide variety of feature templates to meet diverse needs for media processing such as content capturing, editing, recognition, and moderation.

Uses schemes for flexible service process orchestration and integrates the event callback mechanism, which enables you to keep up to date with the task progress.

Online Education

MPS' powerful video transcoding feature can generate video files of different specifications, so video can be played back on various devices, making it highly suitable for the online education industry.

Strengths:

Provides excellent encoding and compression capabilities to minimize the size of the compressed video while retaining the original video definition in online education scenarios, reducing the storage and bandwidth costs.

Supports adaptive bitrate streaming to output multiple bitstreams for one input file. This allows students using different devices to play back video smoothly based on their different network conditions.

Radio and TV Industry

MPS features fast and stable multipart transcoding and supports concurrent task execution and dynamic scaling, satisfying the needs for efficient transcoding in the radio and TV industry.

Strengths:

Supports accelerated transcoding and distributed transcoding to deliver up to 30 times faster transcoding for long videos, allowing radio and TV stations to release content very quickly.

With powerful content discovery, moderation, and quality assurance capabilities, it recognizes scenes and segments of content while guaranteeing the security and high quality of the video.

MPS can transcode videos to 4K and 8K resolutions to meet the needs for video transcoding to UHD in the radio and TV industry.

OTT Smart TVs

MPS supports video transcoding to 2K and 4K resolutions for smart TVs, so as to provide enhanced video quality for viewers.

Strengths:

Supports content transcoding to 4K and 8K resolutions for large-screen devices such as OTT TV to display UHD content.

Supports adaptive bitrate streaming to allow family members to play back video on different devices under complex network conditions.