

媒体处理 产品简介



腾讯云

【 版权声明 】

©2013–2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

产品简介

产品概述

产品功能

产品优势

应用场景

产品简介

产品概述

最近更新时间：2025-03-13 10:30:42

媒体处理（Media Processing Service，MPS）是智能、强大、全面的多媒体数据处理服务，行业支持最全面的音视频编码标准，基于自研编码内核和 AI 算法，提供音视频转码和增强、媒体智能及质检评测等能力。MPS 帮助您提高媒体质量、降低成本，并满足各类场景的音视频处理需求。

产品架构

您可以通过控制台、SDK 或 API 上传的方式，将视频源文件上传至对象存储 COS 的存储桶中，并通过媒体处理的服务编排机制触发媒体处理任务的自动执行。此外，支持使用消息队列 CMQ 接收媒体处理的事件通知，使您能够及时了解转码任务的执行动态。

媒体处理的产品架构图如下所示：



产品功能

最近更新时间：2025-12-12 17:31:12

媒体处理（Media Processing Service，MPS）是智能、强大、全面的多媒体数据处理服务，支持全面的音视频编码标准。基于自研编码内核和AI 算法，媒体处理提供音视频转码、增强、媒体智能、质检评测等能力，可以满足各类场景的音视频处理需求，帮助您提升媒体质量、降低成本。

音视频转码



转码是将原始音视频码流转换成另一个音视频码流的过程，是一种离线任务。通过转码，可以改变原始码流的编码格式、分辨率和码率等参数，从而适应不同终端和网络环境的播放。通过转码功能可以实现：

可实现功能	说明
适配更多终端	将原始视频转码成拥有更强的终端适配能力的格式（如 MP4），使视频资源能够在更多设备上播放。
适配不同带宽	将视频转换成流畅、标清、高清及超清等输出，用户可以根据当前网络环境选择合适码率的视频播放。
改善播放效率	转码可以将 MP4 位于尾部的元信息 MOOV 提前到头部，播放器无需下载完整视频即可立即播放。
节省带宽	采用更先进的编码方式（如 H.265/HEVC）转码，在不损失原始画质的情况下显著降低码率，节省播放带宽。

转码的目标规格包含编码格式、分辨率和码率等参数。您可自定义以下转码相关参数。

参数	类型	详细说明
输入格式	封装格式	3GP、AVI、FLV、MP4、M3U8、MPG、ASF、WMV、MKV、MOV、TS、WebM、MXF。
	视频编码格式	AV1、AVS2、H.264/AVC、H.264 intra、H.263、H.263+、H.265/HEVC、MV-HEVC、H.266/VVC、MPEG-1、MPEG-2、MPEG-4、MJPEG、VP8、VP9、RealVideo、Windows Media Video、WMV2、Quicktime。
	音频编码格式	AAC、ADPCM、AMR、DSD、MP1、MP2、MP3、PCM、RealAudio、Windows Media Audio、VORBIS、AC-3。
输出格式	封装格式	视频：FLV、MP4、HLS (m3u8+ts) 、MXF。
		音频：MP3、MP4、OGG、FLAC、m4a、PCM。
		图片：GIF、WEBP。
	视频编码格式	AV1、H.264/AVC、H.264 intra、H.265/HEVC、MV-HEVC、H.266/VVC、MPEG-2、MPEG-4、MJPEG、VP8、VP9、RealVideo、Windows Media Video、WMV2。
	音频编码格式	MP3、AAC、FLAC、MP2、VORBIS。
封装	删除视频流	如果开启“删除视频流”，转码出来的视频将不包含视频流（仅保留音频流）。
	删除音频流	如果开启“删除音频流”，转码出来的视频将不包含音频流（仅保留视频流）。

音视频转码各类型简介参考下表：

适用输入源	转码类型	描述
视频	普通视频转码	可以降低视频码率、改变原始码流的编码标准、分辨率和帧率等参数，从而适应不同终端和网络环境的播放。
	极速高清转码	极速高清转码是普通视频转码的“升级”功能。能够对不同类型的视频做自适应优化处理，以更低的带宽给用户提供更高清的观看体验。在保证画质甚至提升画质的前提下，节省50%+带宽成本。
	视频自适应码流	可以将输入的源视频文件转成多路适合不同场景下播放的码流，让用户可根据网络选择合适码率的视频播放，提升用户体验。

	转封装	可以改变源视频文件的封装格式。
音频	普通音频转码	可以降低音频码率、改变编码标准、采样率和声道等参数，从而适应不同终端和网络环境的播放。
	音频极速高清转码	音频极速高清转码是普通音频转码的“升级”功能。能够对不同类型的音频做自适应优化处理，以更低的带宽给用户提供更优质的音效体验。
	音频自适应码流	可以将输入的源音频文件转成多路适合不同场景下播放的码流，从而适配不同网络条件，提升用户体验。

音视频增强

基于 MPS 业界领先的音视频 AI 处理模型和丰富的业务数据积累，音视频增强功能可以大幅提升音频和视频质量，广泛应用在 OTT、电商、赛事等场景，带来QoE、QoS 提升等业务收益。

画质增强

统一模型

早退策略

动态调节

色彩增强

音频增强

人脸增强

画面修复

原视频

现视频

去除块状

去除伪影

局部增强

去模糊

视频降噪

细节增强和超分

多分辨率网络模型

多退化低质数据生成

兼顾稳定性和主观判别

视频插帧

高准确性光流运动补偿

支持多种复杂运动

融合前后帧信息

能力丰富

支持灵活

应用广泛

增强类型	能力	说明
视频增强	超分辨率	超分辨率可以识别视频的内容与轮廓，高清重建视频的细节与局部特征，将低分辨率的视频转换为高分辨率的视频，适用于老片修复等场景。
	低光照增强	因环境状况及拍摄摄像头的硬件限制，在某些场景下拍摄的画面存在亮度和对比度缺失问题，导致画面较暗或细节缺失。通过开启低光照增强，可以大幅提升暗区的细节和对比度，提升人眼主观质量。
	HDR	支持 HDR10、HLG，能够获得更宽的色域，展现更多色彩细节，以提供更高品质的视频内容。
	综合增强	通过 AI 的综合分析能力，自动平衡画面中的纹理内容，在去除压缩伪影和毛刺

		的同时增强关键细节，从而提高整个画面的总体主观感受。
	色彩增强	色彩增强可以使画面更接近真实色彩，并在一定程度上进行增强满足人眼的喜好。
	细节增强	细节增强对视频中需要关注的细节（例如赛场草地）进行增强，使画面内容更清晰，内容更丰富。
	人脸增强	通过人脸检测针对视频中人眼视觉特别关注的人脸部分进行增强，使该区域的细节更加清晰，提高主观感受。
	去划痕	去划痕可以修复视频中的划痕和雪花点等破坏的内容。
	去毛刺（去伪影）	由于影片在转码或多次转码过程中对视频进行了多次压缩，会引入块效应、振铃效应、色度渗透和蚊噪等，使得视频画面出现一些影响视觉效果失真，去压缩失真能有效修复编码引入的失真。
	字体增强	对画面中的文字内容进行处理，使文字边界更清晰，提高文字内容的可读性。
	视频降噪	由于影片拍摄中会因为相机和环境引入随机噪点，这里提供降噪服务，在保持细节不损失的情况下，消除画面中的随机噪声。
音频增强	音频降噪	音频智能降噪，去除设备底噪、环境噪声等，适用于录制课程、户外拍摄后期等场景。
	音频分离	分离音视频文件中人声与背景声，或者演唱声和伴奏声，生成独立音频素材，便于后期实现其他风格艺术加工。
	音量均衡	1. 响度标准化：使整体响度水平保持一致，播放时听起来具有相似的音量，避免出现过于响亮或过于安静的问题，提供更好的听觉体验。 2. 音量突变抑制：平滑过于响亮的音频片段，避免出现音量突变的问题，提供更平稳的听觉体验。
	音频美化	1. 杂音去除：减少音频中包含的不需要的杂音或干扰声，提高音频的质量和清晰度。 2. 齿音压制：齿音是指音频中尖锐的、刺耳的声音，通常可能由声音源接近麦克风时产生。压制齿音旨在减少或消除这种不自然的声音，从而改善音频质量。

水印

添加水印是在视频转码或截图时，将特定的图片附加在画面指定位置的过程，是一种离线任务。媒体处理支持以下类型的水印：

- **静态图片水印**：PNG 格式的图片水印，可以是版权方的 LOGO、台标等，常用于表明视频的版权归属。
- **动态图片水印**：APNG 格式的动态图片水印，可以实现水印图片动态变化的效果。

媒体处理支持为视频或截图添加多个水印，并可以指定各个水印在画面中的大小和位置。水印的目标规格包含水印类型、宽高和位置等参数。您可自定义以下水印相关参数。

参数	说明
水印类型（Type）	支持静态图片水印和动态图片水印
水印位置（Position）	水印在视频画面中的相对位置
图片大小（ImageSize）	图片水印占视频画面的大小
图片内容（ImageContent）	图片水印中图片的二进制内容

视频截图

截图是截取视频特定位置的图像并生成图片的过程，是一种离线任务。媒体处理提供以下类型的截图：

- **指定时间点截图：**指定一组时间点，截取视频在这些时间点的图像。
- **采样截图：**按相同的时间间隔对视频截取多张图。
- **雪碧图：**按相同的时间间隔对视频截取多张小图，然后组装成若干大图（即 [雪碧图](#)）。

截图的目标规格，包含了截图文件格式、截图宽高等参数。您可自定义以下截图相关参数。

时间点截图

参数	说明
格式（Format）	截图文件的输出格式，目前仅支持 JPG
宽度（Width）	截图宽度，范围是128px – 4096px
高度（Height）	截图高度，范围是128px – 4096px
填充方式（FillType）	当截图的宽高比与原始视频的宽高比不一致时，对截图的处理方式，即为“填充”。一般有以下几种填充方式： ● 拉伸：对图片进行拉伸，填满整个图片，可能导致图片被“压扁”或者“拉长” ● 留黑：保持图片宽高比不变，边缘剩余部分使用黑色填充 ● 留白：保持图片宽高比不变，边缘剩余部分使用白色填充 ● 高斯模糊：保持图片宽高比不变，边缘剩余部分使用高斯模糊化后填充

采样截图

参数	说明
格式（Format）	截图文件的输出格式，目前仅支持 JPG
宽度（Width）	截图宽度，范围是128px – 4096px
高度（Height）	截图高度，范围是128px – 4096px
采样方式（SampleType）	采样方式分为两种： - 按百分比采样：例如按照5%为间隔采样，生成截图张数将为20张 - 按时间间隔采样：例如按照10s为间隔采样，截图张数取决于视频的时长
采样间隔（Interval）	采样的间隔长度： - 如果按百分比采样，间隔是百分比 - 如果按时间间隔采样，间隔是多少秒
填充方式（FillType）	当截图的宽高比与原始视频的宽高比不一致时，对截图的处理方式，即为“填充”。一般有以下几种填充方式： - 拉伸：对图片进行拉伸，填满整个图片，可能导致图片被“压扁”或者“拉长” - 留黑：保持图片宽高比不变，边缘剩余部分使用黑色填充 - 留白：保持图片宽高比不变，边缘剩余部分使用白色填充 - 高斯模糊：保持图片宽高比不变，边缘剩余部分使用高斯模糊化后填充

雪碧图

参数	说明
格式（Format）	雪碧图文件的输出格式，目前仅支持 JPG
小图宽度（Width）	雪碧图中小图的宽度
小图高度（Height）	雪碧图中小图的高度
小图行数（Rows）	一张大图中有多少行小图
小图列数（Columns）	一张大图中有多少列小图
采样方式（SampleType）	小图采样方式，目前仅支持按照时间间隔采样
采样间隔（Interval）	小图采样的间隔，即隔多久采样一张小图

⚠ 注意

- Width × Columns 需要在128px – 4096px之间（即大图宽度在128px – 4096px之间）。
- Height × Rows 需要在128px – 4096px之间（即大图高度在128px – 4096px之间）。

视频转动图

转动图是选取视频片段生成动图（ GIF 和 WEBP 等）的过程，是一种离线任务。动图是一组连续帧的无缝循环，以较小的体积实现动画效果。

转动图的目标规格，包含了动图格式、宽高和帧率等参数。您可自定义以下动图相关参数。

参数	说明
格式 (Format)	动图文件的输出格式，目前仅支持 GIF 和 WEBP
宽度 (Width)	动图宽度，范围是128px – 4096px
高度 (Height)	动图高度，范围是128px – 4096px
帧率 (FPS)	支持的帧率范围是1fps – 60fps

媒体 AI / 媒体智能

媒体 AI 包括智能字幕、智能擦除、大模型摘要、智能标签、智能 ROI、智能横转竖、人脸识别、语音/文本识别、智能分析等功能。在直播、泛娱乐、教育等多个行业有深度积累，能够满足短剧出海、视频会议、在线教育、平台直播等多场景需求，提高内容制作效率。

智能字幕

智能字幕功能可以将离线视频文件或直播流中的语音信息进行实时语音识别，将其转换成字幕并进行多语言翻译，支持中文、英语、日韩语等多种语言。适用于直播实时字幕、视频转译出海等场景。也支持通过文字识别提取视频画面中的原有字幕并生成字幕文件，适用于老片修复、高清重制等场景。该功能还支持配置热词库、术语库以增加识别和翻译的准确率。查看 [智能字幕接入指引](#)。

技术优势

- 全平台支持：支持处理离线文件、直播流、互动音视频。直播实时同传字幕支持稳态、渐变模式，接入门槛低，无需播放端改造。
- 准确率高：大模型处理，支持热词、术语库，准确率行业领先。
- 语种丰富：支持上百种语种，支持多地方言，支持中英文夹杂等混合语种识别。

- 样式自定义：支持将字幕压制至视频，且字幕样式（字体、字号、颜色、背景、位置等）可自定义，支持页面自定义渲染。



智能擦除

智能擦除能够对视频画面中的字幕、人脸和车牌等元素进行模糊、马赛克或无痕化处理，从而便于内容的传播和分享。该功能广泛应用于短剧平台、短视频平台、跨境电商以及自媒体工作室等多个领域。查看 [智能擦除接入指引](#)。

- 高可定制性：支持针对您的具体的视频场景，对模型参数进行定制化调整，从而提高擦除的准确率和处理效果，确保最终呈现的画面更加自然流畅。

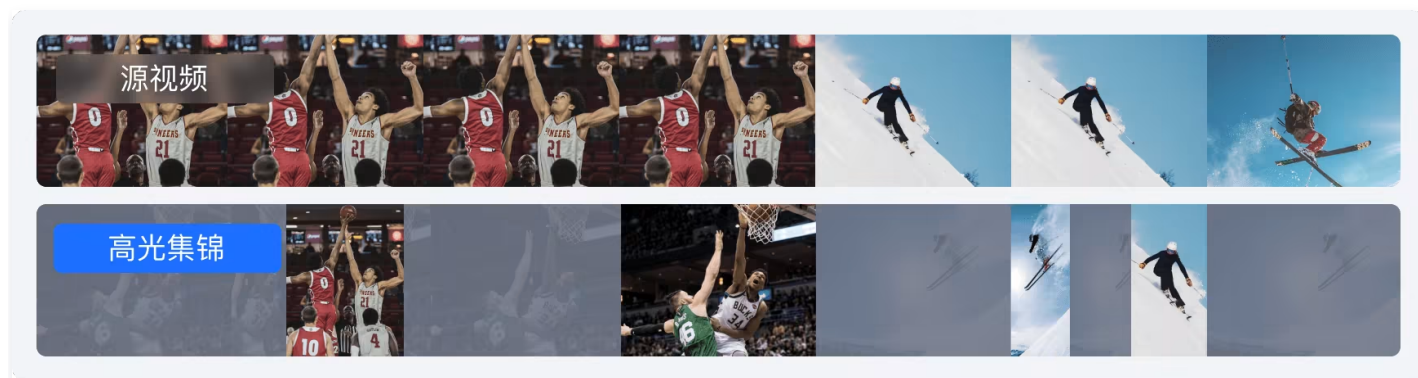


高光集锦 / 精彩集锦

高光集锦定义了精彩场景对应的关键词，例如“男女主角冲突”或“精彩打斗戏”。通过将图像特征和文本特征映射到一个统一的特征空间，我们可以利用这些关键词精准地找到视频中的精彩场景，从而提取出最引人注目的片段。支持影视剧、足球、篮球、绝地求生、王者荣耀等视频场景，支持定制（视频智能服务官方 QQ：303224607，需支付定制费）。查看 [高光集锦接入指引](#)。

技术优势

- 全平台支持：支持处理离线视频，也支持对直播流进行处理，实现对过往的直播内容生成精彩片段。
- 多场景支持：支持影视剧、体育赛事、游戏竞赛、新闻资讯等视频场景。
- 高可定制性：支持针对您的具体的视频场景，对模型参数进行定制化调优。



智能拆条

智能拆条功能可以对视频内容进行结构化分析，将完整的长视频进行分段处理，适用于会议、培训、课程等多种场景。例如，将一段新闻联播的素材拆分成多个独立的新闻事件视频。这可以显著提升新闻和体育类视频的拆条质量，促进二次创作，节省人力和硬件成本。查看 [智能拆条接入指引](#)。

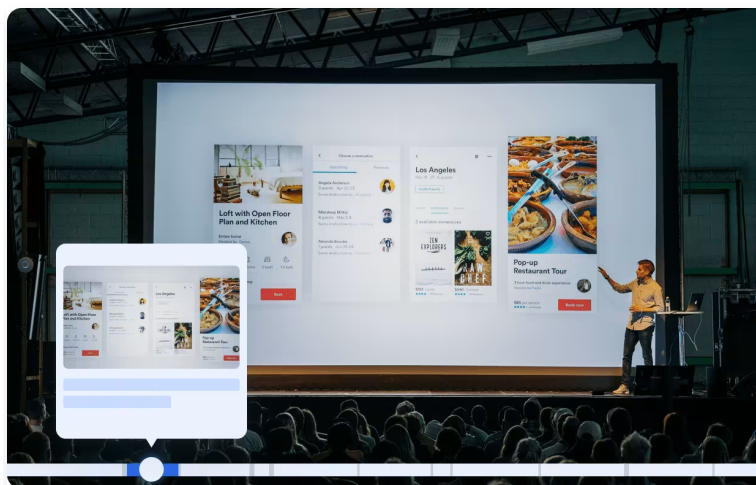


大模型视频摘要

借助 NLP 能力，大模型视频摘要功能可以对教学课程、新闻、演讲等视频场景，基于 ASR 识别结果进行混元大模型摘要提取。其提取和概括力强，无需反复拖拽即可快速理解视频内容。同时还支持离线分段和直播实时分段，自动提取段落摘要及关键词，帮助用户快速定位感兴趣的视频片段，提高使用效率。查看 [大模型视频摘要接入指引](#)。

技术优势

- 全平台支持：支持处理离线视频，也支持对直播流进行处理。
- 多语种支持：支持中文、英文等语种。



00:00

Abstract title

Accounting work has different regulations in different industries. The specific accounting content depends on the special regulations

02:36

Pop-up Restaurant Tour

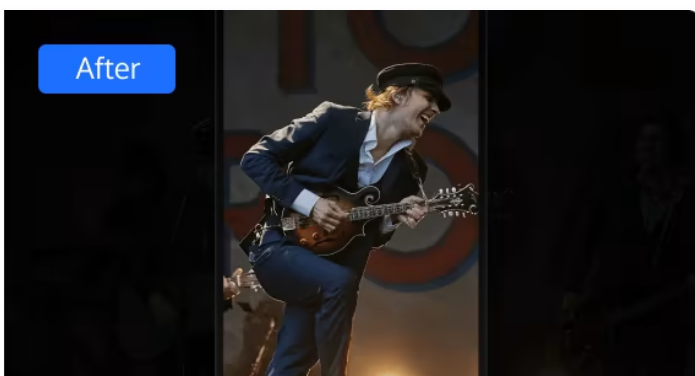
Accounting work has different regulations in different industries. The specific accounting content depends on the special regulations

智能横转竖

横转竖不是单纯的旋转，而是通过识别感兴趣区域（ROI，Region of Interest），将视频裁剪成一定比例的适合移动设备播放的视频。横转竖提供批量生成短视频的能力，能够将现有的横屏视频资源转换成竖屏视频资源。查看[智能横转竖接入指引](#)。

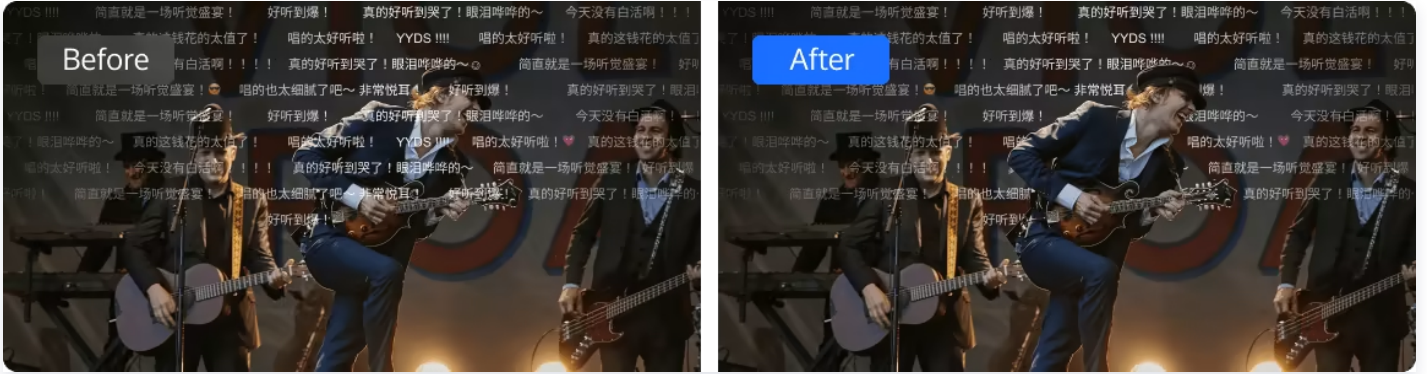
技术优势

- 支持多人物场景：当视频中出现两个人物时，支持自动将画面上下分割，并在竖屏视频中分别显示这两个人物。
- 高可定制性：支持针对您的具体的视频场景，对模型参数进行定制化调优。



ROI 智能识别

ROI 智能识别技术可以实时识别视频中的人脸、游戏人物、主播等关键元素的坐标信息，并将感兴趣区域信息（ROI）随着视频编码传输到播放端。通过播放器结合 ROI 信息，可以实现智能弹幕防遮挡、背景模糊等功能。当前 ROI 智能识别支持处理直播流，如有处理离线视频的需求，请联系我们。查看[直播ROI智能识别接入指引](#)。



智能识别

智能识别基于腾讯实验室的最新研究成果，为您提供视频内容的全方位识别，支持识别视频内的人物、语音、文字以及帧标签，对视频进行多维度结构化分析。

识别类型	功能说明
人脸识别	基于深度学习方案，帮助客户快速识别视频中的人脸信息并快速定位出视频中的人物所在帧画面，以及人脸所在区域。客户可自定义人物库或调用视频 AI 公共人物库进行人脸识别。
语音识别	基于深度学习方案，帮助客户快速识别视频中的声音并转化成文字，支持客户自定义关键词且定位出关键词所在视频的时间点。
文字识别	帮助客户识别视频中出现的文字信息，可用于视频内自定义关键词的提取，也支持竖版文字的识别。
帧标签识别	基于深度学习方案，支持根据客户自定义视频截帧间隔，自动识别截帧画面内的标签，并定位标签所在的视频位置，帧标签涵盖人物、风景、人造物、建筑、动植物、食物等9个大类，包含日常生活的各个信息维度，并且支持在标签体系基础上使用自定义标签，具备迁移学习能力，只需提供原始用户数据便可定制分类器，满足不同类型的用户需求，使标签分类更具灵活性。
片头片尾识别	根据视频画面特征、文字、语音等信息自动识别定位电影/电视剧片头片尾时间点。

智能分析

分析类型	能力说明
分类识别	通过分析视频内容，给视频推荐一个类别。目前支持美食、旅游、动漫、音乐等19大类（支持定制，需支付定制费）。
视频标签识别	基于腾讯深度学习方案，智能识别出最符合视频内容的前五项标签摘要，用于视频推荐、检索等场景，用户可在接口中自行选择返回的标签个数。
智能封面	结合视频画面纹理、场景识别等特征信息自动生成文件封面，支持静态封面输出，提升视频封面体验和点击转化率。

智能审核

智能审核借助 AI 对视频内容（画面、音频及文字三种维度）进行涉黄、违法违规的检测。

检测类型	检测项说明
视频画面审核	对视频画面做涉黄、违规检测，具体检测项如下： ● 涉黄检测 ○ porn：色情 ○ vulgar：低俗 ○ intimacy：亲密行为 ○ sexy：性感 ● 违法违规检测 ○ guns：武器枪支 ○ bloody：血腥画面 ○ explosion：爆炸火灾 ○ violation_photo：违规图标
音频审核	对音频中的文字进行检测，具体检测项如下： ● 涉黄检测：对音频中的文字做涉黄检查，识别出嫌疑关键词。 ● 违法违规检测：对音频中的文字做违法违规检查，识别出嫌疑关键词。
文字审核	对画面中的文字进行检测，具体检测项如下： ● 涉黄检测：对画面中的文字做涉黄检查，识别出嫌疑关键词。 ● 违法违规检测：对画面中的文字做违法违规检查，识别出嫌疑关键词。

媒体质检

支持进行直播、离线视频中的格式质检、画面内容质检、音频质检，并提供视频整体质量检测评分。

质检类型	检测类型	检测项说明
格式质检	● 离线视频格式质检 ● 直播流格式质检	检测 DTS、PTS 问题、分辨率变化，采样率变化、丢帧、重复帧等格式问题。
无参考评分	视频无参考评分	根据多维度检测标准，对视频质量进行百分制打分。
质量审核	画面质量	支持对视频的画面质量做出检测，具体检测项如下： ● JitterResults：画面抖动。

		• BlurResults: 画面模糊。 • AbnormalLightingResults: 低光、过曝。 • CrashScreenResults: 花屏 • BlackWhiteEdgeResults: 画面黑边、白边、黑屏、白屏、纯色屏时间段。 • NoiseResults: 画面有噪点。 • MosaicResults: 画面有马赛克。 • QRCodeResults: 画面有二维码。
	声音质量	支持对视频的声音质量做出检测，具体检测项如下： • VoiceResults: 音频异常，包括静音、低音、爆音。

视频评测

通用全面的视频评测能力，支持对多种视频来源和各类格式的视频进行 VMAF、PSNR、SSIM 和 VMAF-NEG 评分，计算 BD-Rate。支持一站式提供完整的评测报告及指标分析。

分类	功能	说明
离线视频	视频质量评测	添加原视频和对比视频，进行视频质量评测。 • 支持 VMAF、PSNR、SSIM 和 VMAF-NEG 评测方式。 • 可以自定义选取时间段或帧数范围评测。
	BD-Rate 对比评测	选择媒体处理模板，评测不同模板在多种比特率下的视频转码质量差异。 • 支持 VMAF、PSNR、SSIM 和 VMAF-NEG 评测方式。 • 可以自定义选取时间段或帧数范围评测。 • 支持指定码率对比评测分数，或指定 CRF (视频质量分数)对比码率。
直播流	画面质量	实时对比和监控直播流转码前后的画面质量及码率变化。



直播流录制

通过输入直播流地址，可以简单快捷地对直播内容进行录制。支持预设录制周期、单个TS 时长，实现续录功能。

参数	参数说明
单个 TS 时长	单个 TS 时长支持5秒 – 30秒。
录制周期	范围10分钟 – 720分钟，超过设定的录制周期后，将生成新的文件。
续录等待时长	范围60秒 – 1800秒，续录等待时长会直接影响录制文件生成的时间。

终端 SDK

终端 SDK 是腾讯云推出的音视频终端产品能力，包括视频编码、音频增强、视频增强三类功能 SDK，面向不同客户场景需求提供针对性的产品能力。终端 SDK 产品作用于客户终端设备，支持移动端、Web 端、PC 端等多终端接入。具体操作可以参见 [终端 SDK 接入教程](#)。

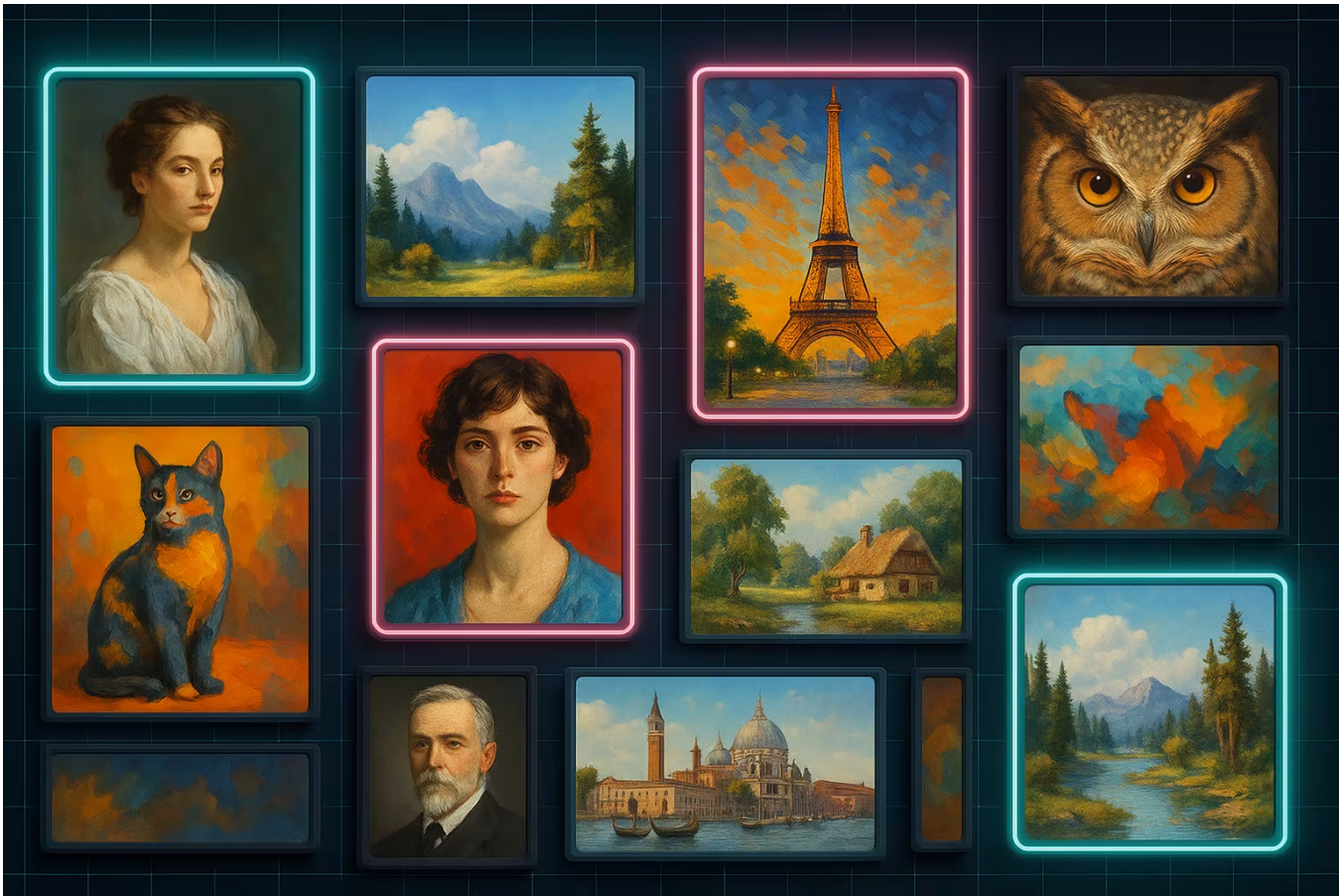
终端 SDK 类型	功能说明
终端视频编码 SDK	腾讯极速高清终端视频编码器是一款针对端侧低算力、低延迟、高画质场景研发的编码器。和硬件编码相比，优势在于： ● 稳定可靠，启动快。 ● 同等画质下节省码率，提升传输的稳定性，节省下行分发带宽，节省存储成本等。

	● 同等码率下提升画质，提升用户体验。 ● 丰富的功能，满足业务多样化需求，如使用 ROI 编码提升人脸区域画质，动态调整编码配置适应网络波动等。
终端音频 SDK	终端音频 SDK 包含标准版、专业版和旗舰版，支持如下功能： ● 声学回声消除 ● 自动增益控制 ● 自适应噪声抑制 ● 回声消除音乐模式 ● 音量均衡 ● AI 智能降噪 ● 音频编码 ● AI Codec
终端增强 SDK	客户端增强 SDK，基于高效的图像处理算法和 AI 模型推理能力，实现终端视频超分辨率、画质增强、插帧等功能，包含标准版、专业版和旗舰版，支持如下功能： ● 标准超分辨率/专业超分辨率/标准超分+增强参数 ● AI 画质增强 ● AI 插帧增强

AIGC 图片和视频生成

腾讯云媒体处理汇集了混元大模型和行业内丰富的 AI 模型，可以一站式调用各类模型，生成图片和视频。结合腾讯云音视频全面的转码、超分、插帧、增强、AI 配音、翻译、字幕、水印等能力，实现一站式的内容制作及媒体处理。

生成类型	功能说明
AI 图片生成	图片生成和自然语言图片编辑能力，支持生成4K图片。
AI 视频生成	支持文生视频、图生视频，稳定保持图片文字、人像和商品一致性。结合视频超分和插帧能力可生成4K 60FPS 的 AI 视频内容。



产品优势

最近更新时间：2025-11-26 11:03:51

高可编排

支持自定义服务流程，可对海量媒体文件进行流程化处理，一站式完成质检、增强、转码、截图、水印等处理环节。还可以为处理流程设置条件，实现智能流转和按需处理。集成事件回调机制，及时掌握任务进度。

极速高清

支持 VP9、H.264、H.265、AV1、AVS3、H.266等编码标准，在保证视频主观质量的前提下节省50%带宽成本。同时支持 4K、8K 实时编码，提供超高清流畅的视频体验。

倍速转码

全球化 Region 部署，自动扩缩容，灵活应对高并发转码需求。长视频支持最高30倍速分布式转码加速，满足极速转码发布需求。

画质重生

业界领先的视频 AI 技术，支持老片修复、标准转高清、高清转 4K 的能力，能够大幅去除视频噪声、毛刺、划痕，提升视频清晰度和色彩丰富度。

功能丰富

支持音视频转码、音视频增强、质检评测、媒体智能、视频截图、内容理解、审核、智能编辑等能力。开放丰富的模板配置能力，支持用户根据自身需求自定义配置。

AI 整合

结合腾讯云先进的 AI 技术，针对海量视频内容提供智能标签、精彩集锦剪辑、智能封面图、人脸识别、语音/文字识别、视频翻译、安全/质量审核等智能服务。

应用场景

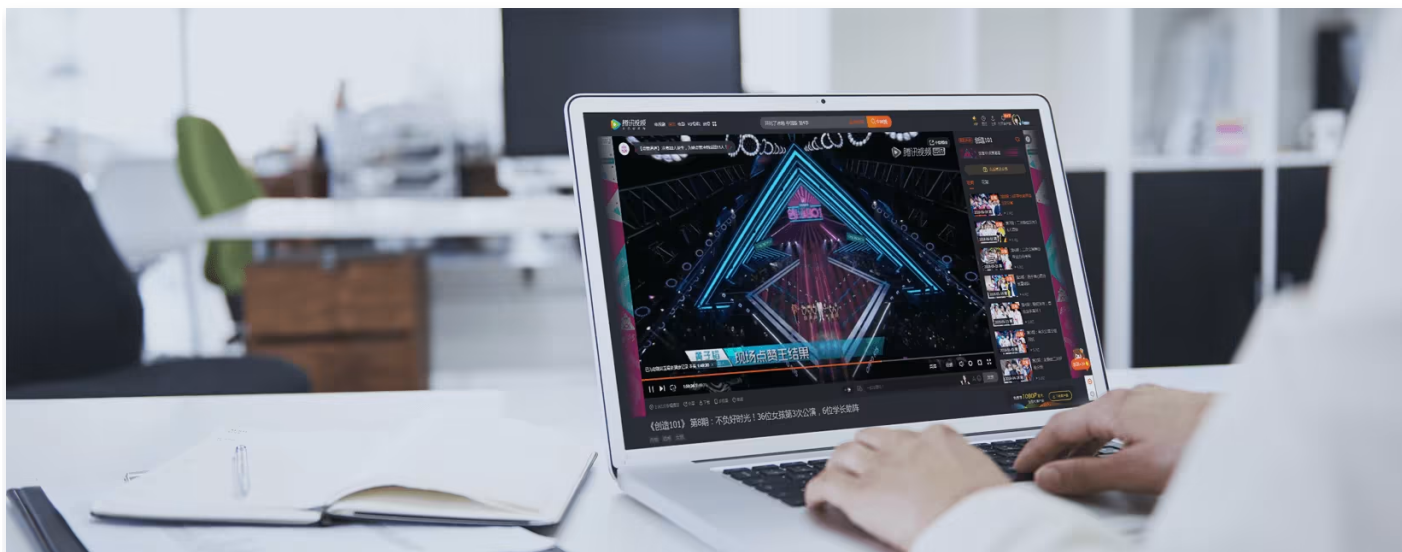
最近更新时间：2025-02-27 17:11:22

视频网站

媒体处理覆盖了主流格式，支持多种分辨率和码率，解决视频网站需要保证不同带宽用户的视频观看体验的需求。

我们能提供：

- 极致的编码体验，支持极速高清转码，大幅降低存储带宽成本。
- 自适应码流转码，满足各种终端、各种网络环境下播放需求。
- 丰富的功能模板、满足内容采、编、识、审等各类媒体处理需求。
- 灵活编排服务流程，集成事件回调机制，实时掌握任务进度。



在线教育

教育客户广泛使用媒体处理能力，来提升媒体资源的质量，并获得更高效的教学体验。

我们能提供：

- 极致的编码压缩能力，针对在线教育场景能够在保持原视频清晰度的前提下极大的压缩视频体积，节约存储及带宽成本。
- 自适应码流转码，支持一进多出生成多路码流，满足多个终端、不同网络条件下的播放需求。
- 终端视频编码和音频降噪能力，能够提升教师在网课中的音视频质量。
- 使用原生录制功能，可以将线上课课堂的学生评论，弹幕和板书信息通过云端机器自动录制下来，完整还原学生端画面。
- 智能字幕能够将老师的授课内容显示为实时字幕，配合大模型摘要功能，在课程结束后生成课程摘要和知识点总结。



短剧译制

在短剧国际化过程中，大量短剧内容需要进行台词和字幕的翻译。通过灵活组合多种媒体处理技术，我们可以快速将视频翻译成多种语言内容。

我们能提供：

- 智能识别-语音识别：通过音频识别技术，自动将剧中人物的台词转换为文本信息。
- 智能识别-文字识别：能够识别视频画面中的字幕信息，并生成相应的字幕文本。
- 字幕擦除：能够无痕擦除视频中的原字幕，使替换不同语种的文字变得更加方便。



广电行业

媒体处理具备高速稳定的分片转码系统，支持多任务并发进行和动态扩容，满足广电行业对转码效率的需求。

我们能提供：

- 支持倍速转码，长视频支持最高30倍速分布式转码，满足广电场景下内容极速发布的需求。
- 强大的内容理解、审核、质检能力，在保证视频安全、高质的前提下能够对内容进行场景及片段识别，满足广电场景下对于内容运营的需求。
- 支持4K、8K超高清转码，满足广电行业超高清内容转码需求。

- 通过媒资质检能力，识别点播和直播内容的格式问题，以及画面/声音内容问题，有效的对媒资内容进行准入质检，和质量监控。



OTT 智能电视

媒体处理支持各类转码格式和高分辨率，高帧率的视频内容处理，满足智能电视的超清需求，从而为客户提供更为细腻真实的视觉效果。

我们能提供：

- 支持4K/8K超高清内容转码，满足 OTT 电视等大屏设备对于超高清内容的使用需求。
- 自适应码流转码，满足终端设备在复杂的家庭网络环境下的播放需求。
- 通过视频超分能力，可以在服务端和终端设备上，将画面内容进行超分处理，为最终用户提供更高分辨率的画面效果。



AIGC 内容质量提升

在图像和视频大模型训练和生成场景，也融合了腾讯云音视频丰富的技术能力。

我们能提供：

- 在视频训练数据清洗环节，通过媒体质检能力，可以将存在格式问题和画面问题的训练素材剔除，提升训练视频的整理质量。
- 在视频生成场景，视频超分能力可以将 AI 生成的视频清晰度进一步提升，并可以通过智能补帧能力获得更加流畅的画面效果。
- 为大模型生成结果添加浮现式水印和隐藏式水印。

