

# 媒体处理

## 产品简介

### 产品文档



腾讯云

**【版权声明】**

©2013-2019 腾讯云版权所有

本文档著作权归腾讯云单独所有，未经腾讯云事先书面许可，任何主体不得以任何形式复制、修改、抄袭、传播全部或部分本文档内容。

**【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。

**【服务声明】**

本文档意在向客户介绍腾讯云全部或部分产品、服务的当时的整体概况，部分产品、服务的内容可能有所调整。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模式的承诺或保证。

## 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

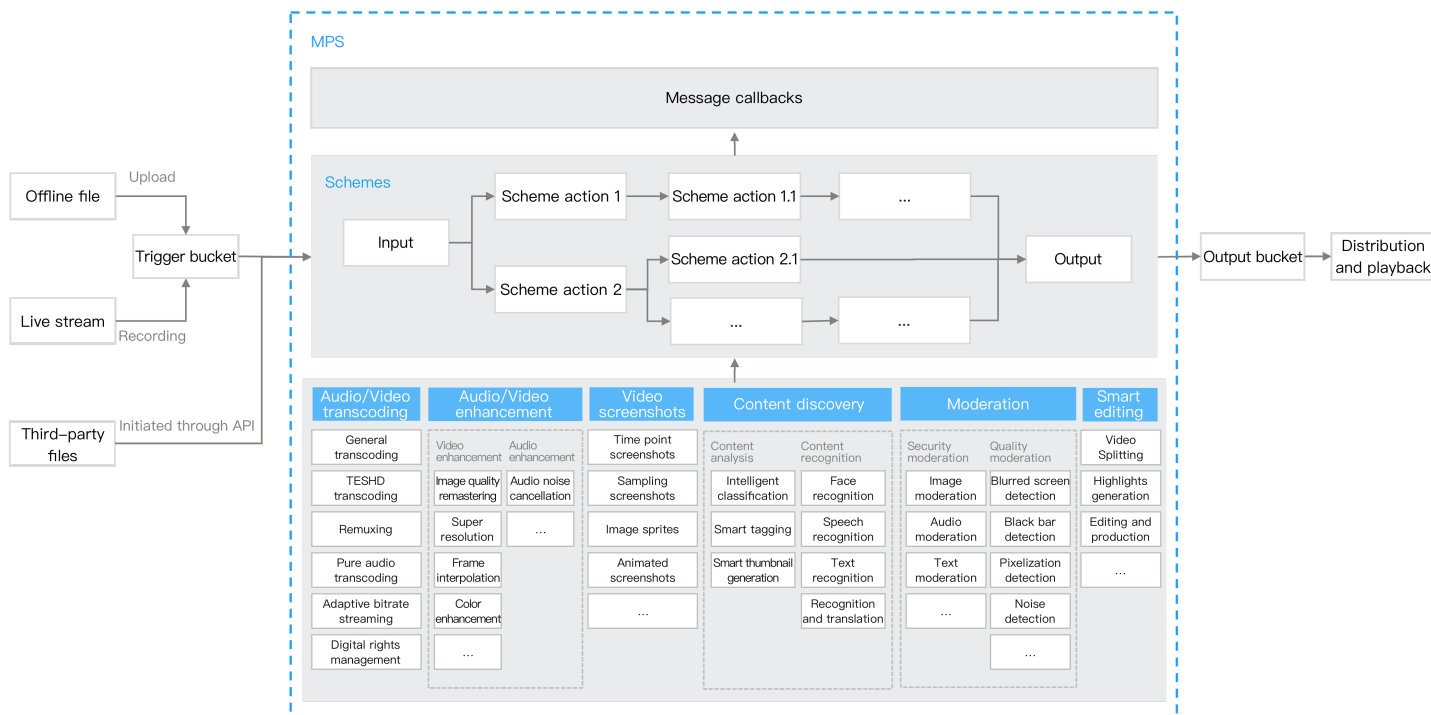
## 产品概述

最近更新时间：2022-08-10 15:36:52

媒体处理（Media Processing Service, MPS）是一种云端音视频处理服务。基于腾讯多年音视频领域的深耕，为您提供极致的编码能力，大幅节约存储及带宽成本、实现全平台播放，同时提供视频截图、音视频增强、内容理解、内容审核等能力，满足您在各种场景下对视频的处理需求。

## 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，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，且支持消息队列 CMQ 接收媒体处理的事件通知，使您可以及时掌握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：2022-08-10 15:36:52

媒体处理将音视频文件转码为不同码率和分辨率的格式，以满足不同网络带宽和终端设备的用户需求。支持以下功能：

## 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）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.263、H.263+、H.265、MPEG-1、MPEG-2、MPEG-4、MJPEG、VP8、VP9、RealVideo、Windows Media Video、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

参数	类型	详细说明
		图片：GIF、WEBP
	视频编码格式	H.264/AVC、H.265/HEVC、AV1
	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）

## 音视频增强

通过画质修复和画质增强两大模块结合 AI 算法，提升分辨率的同时，能够提供视频去噪、轮廓修复、超分辨重建等功能，适用于 UGC/PGC 视频质量提升，老片翻新，4K 生产等业务场景。

能力	说明
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。
去伪影（毛刺）	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的失真，去压缩失真能有效修复编码引入的失真。
去划痕	由于影片在拍摄、保存和转存过程中一些未知因素导致胶片产生划痕和雪花点等破坏，去划痕可以修复视频中的划痕和雪花点等破坏的内容。
细节增强	由于拍摄相机质量、后期保存和转码过程中导致视频细节不够清晰，细节增强对视频中需要关注的细节进行增强，使画面内容更清晰，内容更丰富。
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。
超分	针对一些影片分辨率较低，不适应如今对高分辨率影片的要求，提供超分辨率能力对影片细节进行增强修复，利用 AI 模型达到高分辨率输出的同时有效细节更清晰。
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。

能力	说明
色彩增强	因拍摄设备的色彩问题或影片的保存问题，导致影片的色彩存在一定的失真或可增强的情况，色彩增强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。
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亮度和对比度缺失，导致画面较暗或画面较暗区域细节看不到，因此暗场景增强能够自动识别场景对画面进行自适应增强，特别在暗场景下能大幅提升暗区的细节和对比度，提升人眼主观质量。
HDR	SDR 转 HDR 能力可将普通 SDR 视频转化为 HDR 视频，色深提升至10bit，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。
插帧	通过对原始视频帧间补充新的视频帧，给用户提供更加流畅丝滑的视觉效果。另外，也能够帮助解决低帧率老视频中卡顿、抖动等画质问题。

## 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： <ul style="list-style-type: none"><li>• 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”</li><li>• 留黑：保持图片宽高比不变，边缘剩余部分使用黑色填充</li><li>• 留白：保持图片宽高比不变，边缘剩余部分使用白色填充</li><li>• 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</li></ul>

## 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： <ul style="list-style-type: none"><li>• 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</li><li>• 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</li></ul>
采样间隔 (Interval)	采样的间隔长度： <ul style="list-style-type: none"><li>• 如果按百分比采样，间隔是百分比</li><li>• 如果按时间间隔采样，间隔是多少秒</li></ul>
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： <ul style="list-style-type: none"><li>• 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”</li><li>• 留黑：保持图片宽高比不变，边缘剩余部分使用黑色填充</li><li>• 留白：保持图片宽高比不变，边缘剩余部分使用白色填充</li><li>• 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</li></ul>



## 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

## 内容理解

### 内容识别

内容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。

## 内容分析

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

智能审核包括安全审核和质量审核两大类。

**安全审核**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。

**质量审核**支持检测直播、点播视频中的画面帧以及声音质量，涵盖花屏、黑边、马赛克、噪声等全方位13项检测类型，并提供主播视频整体质量检测评分。

审核类型	检测类型	检测项说明
安全审核	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： <ul style="list-style-type: none"> <li>• <b>涉黄检测</b> <ul style="list-style-type: none"> <li>◦ porn：色情</li> <li>◦ vulgar：低俗</li> <li>◦ intimacy：亲密行为</li> <li>◦ sexy：性感</li> </ul> </li> <li>• <b>违法违规检测</b> <ul style="list-style-type: none"> <li>◦ guns：武器枪支</li> <li>◦ bloody：血腥画面</li> <li>◦ explosion：爆炸火灾</li> <li>◦ violation_photo：违规图标</li> </ul> </li> </ul>
	音频审核	对音频中的文字进行检测，具体检测项如下： <ul style="list-style-type: none"> <li>• <b>涉黄检测</b>：对音频中的文字做涉黄检查，识别出嫌疑关键词</li> <li>• <b>违法违规检测</b>：对音频中的文字做违法违规检查，识别出嫌疑关键词</li> </ul>
	文字审核	对画面中的文字进行检测，具体检测项如下： <ul style="list-style-type: none"> <li>• <b>涉黄检测</b>：对画面中的文字做涉黄检查，识别出嫌疑关键词</li> <li>• <b>违法违规检测</b>：对画面中的文字做违法违规检查，识别出嫌疑关键词</li> </ul>
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： <ul style="list-style-type: none"> <li>• JitterResults：画面抖动</li> <li>• BlurResults：画面模糊</li> <li>• AbnormalLightingResults：低光、过曝</li> <li>• CrashScreenResults：花屏</li> <li>• BlackWhiteEdgeResults：画面黑边、白边、黑屏、白屏、纯色屏时间段</li> <li>• NoiseResults：画面有噪点</li> <li>• MosaicResults：画面有马赛克</li> <li>• QRCodeResults：画面有二维码</li> </ul>
	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： <ul style="list-style-type: none"> <li>• VoiceResults：音频异常，包括静音、低音、爆音</li> </ul>

## 智能编辑

智能编辑基于 AI 以及腾讯多年的音视频技术，支持多维度全方位理解视频内容，智能集锦以及智能拆条的能力，辅助视频内容生产。

能力类型	能力说明
------	------

能力类型	能力说明
拆条	对视频内容进行结构化分析，根据视频的场景信息、语音信息及文字信息，对视频生成智能拆条，支持新闻拆条、广告拆条。
智能集锦	基于视频时域/空域特征匹配、场景识别、目标检测等技术，自动生成视频精彩片段，支持足球、篮球、绝地求生、王者荣耀等视频场景（支持定制，需支付定制费）。
剪辑制作	支持视频的裁剪、拼接、图片转视频、贴片、帖字、画中画、音频编辑等操作

# 产品优势

最近更新时间：2022-08-10 15:36:52

## 高可编排

支持自定义服务流程，可对海量媒体文件流程化处理，一站式完成转码、截图、水印等基本操作，集成事件回调机制，及时掌握任务进度。

## 极速高清

支持 H.264、H.265、AV1 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，能够大幅祛除视频噪声、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、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、智能封面图、人脸识别、语音/文字识别、智能编辑、安全/质量审核等智能服务。

# 应用场景

最近更新时间：2022-08-10 15:36:52

## 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

媒体处理具备强悍的视频转码功能，可以针对不同的终端生成对应规格的视频，满足在线教育行业多端播放的要求。

**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。



## 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。



## OTT 智能电视

媒体处理支持 2K 和 4K 转码，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。

