

媒体处理 场景实践教程



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

场景实践教程

画质提升场景

音视频成本优化场景

短剧出海场景

生成式场景

在线教育场景

场景实践教程

画质提升场景

最近更新时间：2025-03-11 16:34:52

概述

无论在点播视频场景还是直播场景，画面清晰、细节丰富的视频内容都会带来更好的用户留存和商业转化，因此各大平台都努力优化自己的视频内容。在游戏、秀场、教育、影视等不同的内容类型中，画质提升的目标和技术细节存在较大差异。腾讯云媒体处理凭借自身优秀的技术能力和丰富的行业经验，在画质提升场景中，提供多样化的技术解决方案。

客户核心痛点

- 我所在的行业，画质提升应当使用哪些能力提升画面效果？
- 如何用更高效的方式提升画质，达成提升质量的目标，但不至于承担过高的成本？
- 支付了费用进行增强处理，我应该如何量化画质提升的效果？

我们提供的解决方案

丰富的音视频增强能力

针对不同类型的画质提升诉求，腾讯云媒体处理提供多样化的音视频增强能力，用户既可以使用单项的原子能力，也可将多项能力组合使用达到更好的效果。

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。
	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	● 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。 ● 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	● 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。 ● 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。

增强前后效果参考下图：



在 [媒体处理控制台](#) 上可以创建和管理您的音视频增强模板，每个模板中均预设有各项增强参数，创建模板参数页面也提供了针对不同场景的预设参数。您可以根据自身业务场景快速配置合适的增强参数。

视频增强详细参数配置

查看参数描述

推荐参数

通用场景

短视频场景

高清影视剧场景

秀场/电商/大会/讲座场景

游戏场景

老片修复场景

帧率

与源文件一致

自定义

超分辨率①

低光照增强

HDR①

综合增强

弱

适中

强

色彩增强

细节增强

人脸增强①

去划痕

去毛刺

视频降噪

按需处理方案

在提供多样化的音视频增强能力的同时，腾讯云音视频还提供“质检+转码&增强”的按需处理方案。针对以优化成本为目标的使用场景，在执行增强处理前，可以先对视频源进行一次质检处理。再根据质检结果，选择针对性的增强模板。在画质增强处理环节之后，还可以通过媒体质检的无参考评分功能，对比处理前后的视频评分，以量化画质提升的效果。

功能	功能说明	如何接入
媒体质检	媒体质检支持 ● 格式质检：流状态异常，流信息异常，容器封装异常，码流诊断，时间戳异常等。 ● 视频&音频内容质检：色彩失真、低光照、暗角、对比度异常、重影、模糊、马赛克、噪点等。 ● 无参考评分。	参见 媒体质检接入 文档。
音视频增强	在按需处理场景，主要针对质检中暴露的音视频问题，进行针对性的处理。包括但不限于： ● 低光照增强 ● 综合增强 ● 人脸增强 ● 音频降噪 对于不存在问题的视频则无需增强处理，该方式可以显著节省处理成本。	参见 发起媒体处理 文档。

音视频成本优化场景

最近更新时间：2025-03-11 16:34:52

概述

随着直播电商、在线教育、远程会议、短视频、短剧、AI 创作等行业趋势的发展，各大平台上视频类内容的体量快速增长。服务提供商面临着提供无缝、高效体验的同时，还需控制运营成本的双重挑战。有效的成本优化策略不仅能够确保服务的可持续性，还能提升市场竞争力。腾讯云媒体处理提供多种产品能力，帮助客户降低音视频处理成本，以及存储成本。

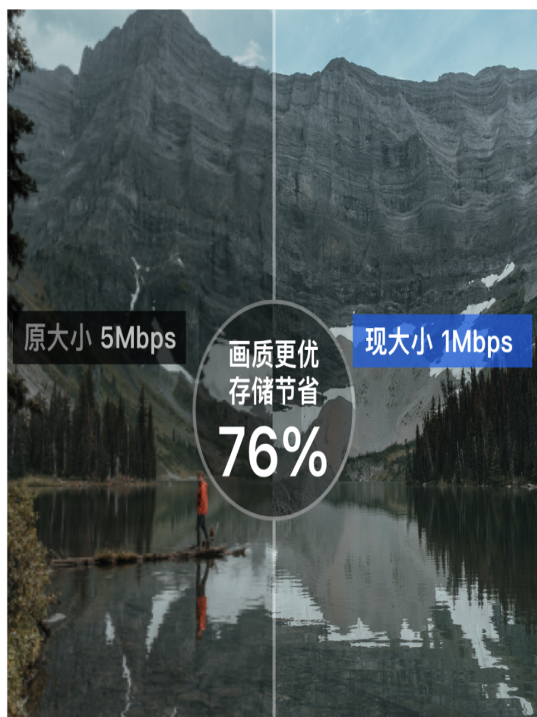
客户核心痛点

- 视频内容增长迅速，存储成本和数据传输费用增加，需要技术手段降低这部分成本。
- 在处理音视频内容时，功能项开启较少会导致优化效果不明显，而同时开启过多功能项则会导致处理成本过高。能否根据不同视频分门别类，或者根据不同视频存在的问题进行差异化处理，以确保成本投入发挥最大效用？

我们提供的解决方案

极致的编码压缩能力

支持 VP8、VP9、H.264、H.265、AV1、AVS3、H.266等编码标准，支持 4K、8K 实时编码，提供超高清流畅的视频体验。在保证视频主观质量的前提下，节省50%带宽成本。最近3年连续获得 MSU 视频编码大赛第一名；Streaming Learning Center 编码测评全项最优。您可以在 [线上页面体验](#) 转码效果。



极速高清转码，大幅降低存储带宽成本



在 [媒体处理控制台](#) 上已为您提供了多种音视频转码模板，每个模板中均预设各项转码参数，您可以根据自身业务场景快速选取合适的转码模板。

按场景筛选

适用视频场景

通用

通用转码场景

指定场景

PGC高清影视

- 保证观看体验

高清素材

- 维持较高画质

UGC内容

- 提升QOS/QOE指标

秀场/电商类

- 强调细节、ROI区域

教育类

- 强调文字、图像

转码策略

极致压缩

极致压缩 推荐

标准压缩

高压压缩-码率优先

中压缩-综合最优

低压压缩-画质优先

同时您也可以创建新的模板，来自定义配置各项参数。创建模板和配置各项参数，具体可参见 [音视频转码模板](#)。关于各项音视频转码的费用详情，可以前往 [计费说明](#) 了解。

按需处理方案

在提供高质量的转码能力的同时，腾讯云音视频还提供“质检+转码&增强”的按需处理方案。针对以成本优化为目标的使用场景，执行转码处理前，可以对视频源进行一次质检处理。再根据质检结果，选择针对性的转码模板。

功能	功能说明	如何接入
媒体质检	媒体质检支持 - 格式质检：流状态异常，流信息异常，容器封装异常，码流诊断，时间戳异常等。 - 视频&音频内容质检：色彩失真、低光照、暗角、对比度异常、重影、模糊、马赛克、噪点等。 - 无参考评分。	参见 媒体质检接入 文档。
普通转码/极速高清转码/自适应码流转码	在按需处理场景，媒资经过媒体质检环节后，根据质检结果自动判断，对存在格式问题、视频码率过高的视频进行转码处理。不存在问题的视频则无需转码，该方式可以显著节省转码成本。	参见 发起媒体处理 文档。

转码后评测效果

在完成转码后，用户还可以通过转码评测功能，定量的衡量转码效果。腾讯云媒体处理提供完整的转码评测能力，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。

分类	功能	说明	使用方法
点播视频	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。	控制台 使用说明 。
	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。	控制台 使用说明 。
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。	内测中，请联系商务。



短剧出海场景

最近更新时间：2025-03-12 16:49:12

概述

随着全球化的加速，短剧作为一种轻量级的文化产品，越来越多地受到国际市场的欢迎。为了帮助内容创作者和制作人将短剧成功地推向海外市场，腾讯云媒体处理产品提供了一套全面的短剧出海解决方案。本文档旨在介绍该解决方案的核心功能和使用方法，帮助用户快速实现短剧的国际化和本地化。



左：原视频中文字幕中文语音；右：处理后生成多语言字幕和语音；注：图片由腾讯元宝 AI 生成

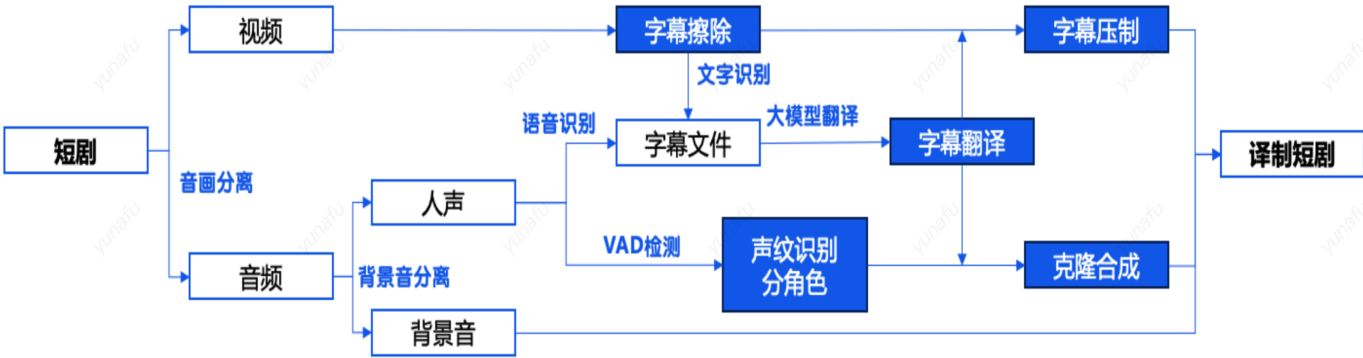
短剧出海核心需求

- 海量视频在不同市场推广，人工处理效率低下，如何批量处理？
- 如何自动从原视频中识别语音内容，智能一键生成准确的多语言字幕，帮助非中文观众理解短剧？
- 如何才能擦除或替换视频中的品牌 Logo、水印和原字幕内容？
- 如何批量将视频的原语言对白替换成目标市场的语言，并重新配音，以提升本地观众的观看体验？

我们提供的解决方案

国产短剧出海急需解决字幕翻译、配音转换、多语种翻译等问题，我们提供短剧译制一站式方案。

短剧翻译处理流程



功能	功能说明
一站式短剧译制（去字幕+语音字幕&翻译+音频合成&替换）	该功能可一站式完成原始字幕擦除、字幕&语音翻译、多语言字幕&语音替换。使用该场景能力，将对输入视频按以下步骤依次处理： 1. 识别原始视频画面中的字幕，然后进行擦除。 2. 识别原始视频的音频，生成字幕文件，并对其进行翻译，再对翻译后的字幕： 2.1. 渲染到原视频画面字幕位置。 2.2. 通过语音合成，替换原有视频的音轨。 3. 合成新的视频。 效果示例  左图：原始中文短视频；右图：处理后输出视频具有英文字幕、英文配音。

不同原子能力也可以单独接入，相关介绍如下：

原子能力	功能说明
水印、Logo、字幕等内容擦除	去除视频中的水印、Logo、字幕等内容。也支持指定区域进行擦除。支持模糊、无痕化等处理效果。 目前支持十多种常见水印的智能擦除。对于不在支持范围内的水印，也支持定制化训练服务，但将单独收取模型训练费用。
人脸车牌敏感信息模糊/马赛克	对视频画面中的人脸、车牌等敏感信息进行模糊或马赛克处理。
人声背景声分离	分离音频中的人声和背景声。
提取字幕（基于文本识别）	基于视频画面中的原有字幕自动提取并生成字幕文件。支持配置热词库，提高识别准确率。
生成字幕（基于语音识别）	基于视频原有音频自动生成字幕文件。支持配置热词库，提高识别准确率。
字幕翻译	翻译原字幕，生成多语言字幕，以满足不同语言背景观众的需求。支持配置术语库，提高翻译准确率。
字幕压制至视频	基于自动生成的字幕文件，通过转码压制到视频画面，样式、字体、字号、字体颜色、背景色均可自定义压制。
文字合成音频	通过提取出的字幕文件，合成新的多语言音频文件。
替换原视频音频	将生成的多语言音频，替换掉原视频音频，以提升本地观众的观看体验。
数字版权保护 DRM	对于有版权的内容，支持商业机 DRM 及 HLS 私有加密。目前支持 FairPlay 和 Widevine 的 DRM 加密格式。

成功案例

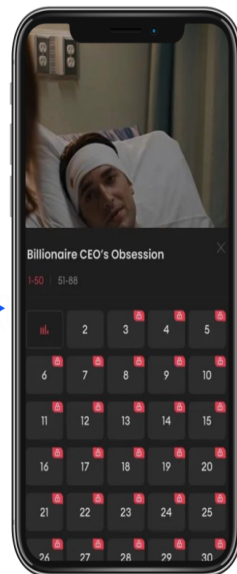
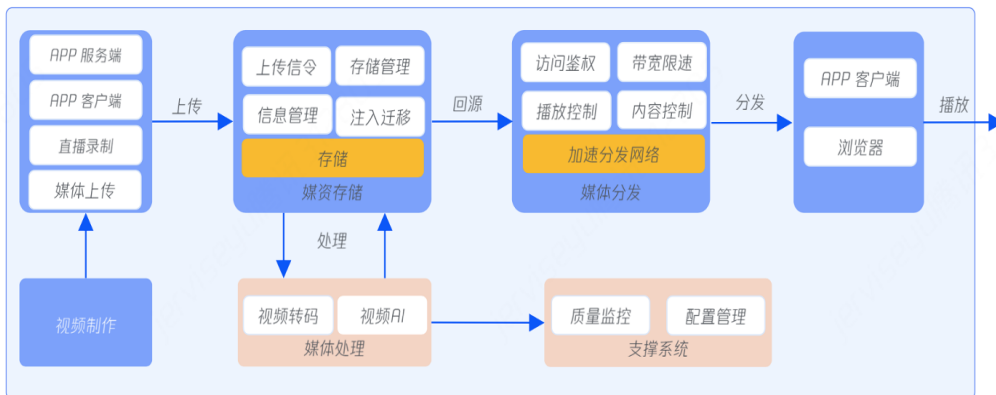
某头部短剧平台客户，在产品链路多个环节使用腾讯云音视频产品。媒体处理视频转码、视频 AI 功能帮助客户显著提升视频播放清晰度和节省存储成本。背景及方案如下图所示：

客户情况

客户诉求： 客户自建短剧视频播放APP，目前已在海外20+国家/地区上线，月均用户量千万+。
客户对视频播放清晰度、稳定性、流畅体验要求非常高，同时客户自制的短剧在分发过程中需加密保障不被盗用、盗播。

产品方案： QUIC上传、极速高清、智能降冷、媒体处理、媒体分发、DRM加密

解决方案



生成式场景

最近更新时间：2025-03-11 16:34:52

概述

大模型音视频内容生成，需要海量的高质量训练素材。当前的行业竞争趋势也向着提供更高清晰度，更流畅，更真实的 AIGC 内容发展。在该场景下，腾讯云音视频将自身质检、音视频增强、画面超分技术，与生成式场景深入结合，极大提升了生成结果的质量。

生成式场景核心需求

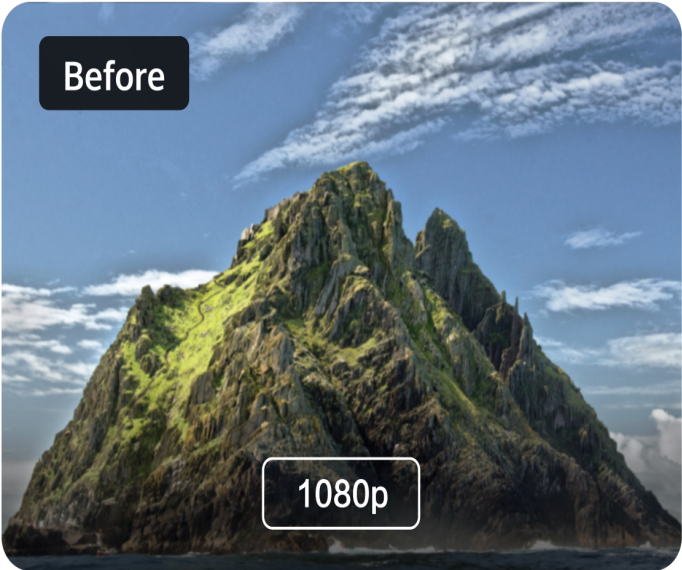
- 在视频训练素材收集和清洗阶段，需要能够甄别质量不佳的视频信息，将问题视频分拣。
- 需要提升生成视频的分辨率，帧率，色彩效果，让画面更清晰流畅。
- 对生成的结果添加必要的水印信息，满足信息溯源和合规要求。

我们提供的解决方案

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。

- 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。

Before



1080p

After



2k 超清

功能	功能说明	如何接入
----	------	------

媒体质检	媒体质检支持： ● 格式质检 ● 视频&音频内容质检 ● 无参考评分 该功能在模型训练素材清洗阶段，可以检测视频格式问题，并可以识别画面问题如色彩失真、低光照、暗角、对比度异常、重影、马赛克等问题。通过质检剔除问题视频后，整体训练素材的内容质量会显著提升。	参见 媒体质检接入文档 。
音视频增强	音视频增强功能，可以提供视频去模糊、降噪、细节增强、文字增强、色彩增强、人脸增强、视频超分、视频插帧等能力。在大模型视频生成场景，能够将原始生成的720P或者1080P分辨率的视频，超分成2K、4K分辨率，并可以通过插帧、细节增强、色彩增强等功能进一步提升画面质量。	参见 发起媒体处理文档 。
视频水印	支持为视频添加浮现式水印（明水印）和隐藏式水印（暗水印），满足信息溯源和合规要求。 支持自定义水印样式如水印图片、水印位置、水印大小等。	

在线教育场景

最近更新时间：2025-04-08 10:31:02

概述

线上教育行业正经历着技术驱动的深刻变革，随着用户规模的增长和新技术的发展，如何在有限带宽下保障高清教学画质、降低运营成本，有效融合创新性技术成为行业核心痛点。本文档旨在介绍腾讯云媒体处理在教育场景的核心功能和使用方法，帮助用户实现技术落地和降本增效。

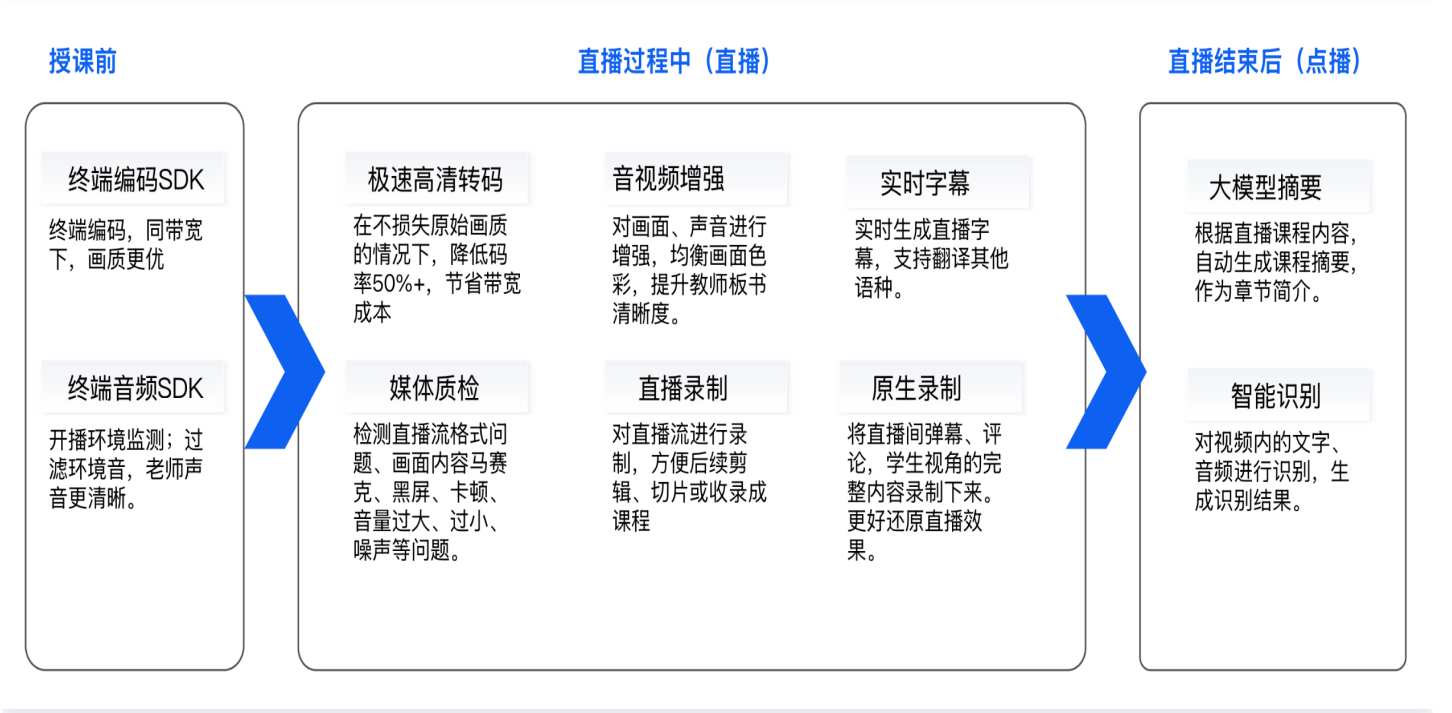
使用场景

清晰稳定的直播效果

- 需求描述：需要画面更清晰，让学生看清板书和课件信息；声音更稳定清楚，确保学生能够听清内容。
- 解决方案：通过终端视频编码和音频降噪能力，提升教师在网课中的音视频质量，确保画面清晰、声音稳定；使用极速高清转码和快直播能力，提供低延迟高画质的直播效果。

AI 能力提升教学水平

- 需求描述：需要在课堂场景以及课后场景应用 AI 能力，协助提升教学水平。
- 解决方案：使用智能字幕功能，将老师的授课内容显示为实时字幕，并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，提升教学效果。



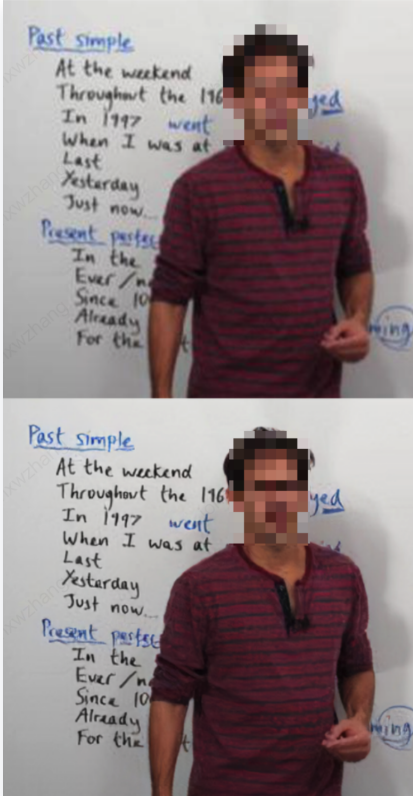
提升教师端开播效果

通过提升教师端的编码效果，和音频降噪能力，可以优化老师在开播端推流的音视频质量。

功能	功能说明	如何接入
终端编码 SDK	终端编码 SDK 集成在教师开播的客户端，通过开播侧的软编编码，可以获得更高质量的视频压缩效果。和硬件编码相比，优势在于： 1. 稳定可靠，启动快。 2. 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。 3. 同等码率下提升画质，提升用户体验。 4. 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。	步骤一：使用前需要在控制台 开通测试 License 或 购买正式 SDK License； 步骤二：技术接入请参考 极速高清终端编码 SDK 接入 流程。
终端音频 SDK	终端音频 SDK 集成在客户端，可以显著提升音频质量、消除回声及噪声。其技术使用到了： 1. 自适应噪声抑制，消除语音信号中的背景噪声成分、降低干扰，以提高语音感知质量。 2. 声学回声消除，主要解决音频通信过程中的回声问题。 3. 自动增益控制，自动识别音频流响度，动态调节音量水平，使其始终保持在一个舒适的范围之内。 4. 终端音频质检能力可以对音频输入质量进行评分，及时识别音频问题并提示	步骤一：使用前需要在控制台 开通测试 License 或 购买正式 SDK License； 步骤二：技术接入请参考 极速高清终端音频 SDK 接入 流程。

提升学生端的观看效果

功能	功能说明	如何接入
极速高清	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。	参见 发起媒体处理 文档。
媒体质检	媒体质检支持： ● 格式质检 ● 视频&音频内容质检 ● 无参考评分 该功能在直播场景可以检测直播画面和格式问题，主动告警和上报，协助运营人员及时发现直播异常。	参见 媒体质检接入 文档。

	在点播场景可以对音视频内容进行检测，为视频内容打分，并可以通过串连转码、极速高清、音视频增强功能，来实现针对性的问题修复。	
音视频增强	业界领先的音视频 AI 增强算法，丰富的业务场景应用模型，全面提升音频和视频质量。支持分布式实时画质增强，包含视频去毛刺、降噪、色彩增强、细节增强、人脸增强。 在教育场景使用，可以优化视频画面色彩、提升板书文字的清晰度。如下图： 	参见 发起媒体处理 文档。
原生录制	传统的录制功能仅能录制纯净的直播流画面，不能以观众的视角完整的录制客户端上的音视频内容。为实现“所见即所得”，将客户端原生内容作为录制源，可将用户终端的“本地录屏”功能模块迁移至云渲染 Windows 实例或 Android 容器，推流至腾讯云直播后，结合录制模块，生成录制文件。 该功能可以用来生成课程回看和教师课件。	参见 云端原生录制 文档。
智能字幕	智能字幕功能可以将点播视频文件或直播流中的语音信息进行实时语音识别，将其转换成字幕或进行多语言翻译。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。	参见 字幕生成及翻译 文档。

在教育场景智能字幕功能，可以应用在教师上课时在直播间实时生成字幕；也可为点播视频添加字幕，提升用户的观看体验。

大模型摘要可以在智能识别的基础上，为视频内容生成总结摘要。为视频内容自动化的生成知识点、和课程内容总结。

大模型摘要



第02讲 等差数列与等比 数列以及性质

主要讲述了等差数列的相关知识。首先，介绍了等差数列的定义和性质。如递增、递减、常数列等。然后，通过函数和数列的关系，阐述了如何求解等差数列的公差。接着，讲解了如何求和。最后，通过例题，讲解了如何应用等差数列的性质解决实际问题。要透过现象看本质，才能更好地理解和解决题目。

智能总结由机器自动生成，仅供参考

三章 发评论 话题

00:03	07:57	12:43	17:41	24:30	29:40	43:41	49:33	59:29	1:05:15
数列知识点讲解	数列公式推导与	数列公式应用与	等差数列性质及	数学方程求根技	数学题目解析与	等差数列应用题	高考试题解析	高考试题解析与等	数列问题解答与
与高考试题解析	应用	等差数列性质解	快速求和技巧	巧与应用	奇数项求和方法	解析与解答	与技巧	等差数列性质	等差数列理解

- 可通过 [体验馆](#) 进行效果体验。
- 能力接入可参见 [大模型视频摘要接入](#) 文档。