

腾讯同传

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

实时会议

文件转写

历史文件管理

智能优化

产品优势

应用场景

核心技术

产品简介

产品概述

最近更新时间：2024-08-23 17:23:20

腾讯同传能够智能的帮助企业提升日常办公效率和准确度。腾讯同传基于语音识别、语义理解、机器翻译等人工智能技术，结合机构、企业办公应用场景，提供中英文互译、中英文实时转写、会议记录编辑成稿、角色分离、历史文件管理、效果优化等功能。

功能架构



腾讯同传可灵活适应多种语音场景下的使用需求。也提供智能优化功能，用户可根据业务需求，定制识别翻译内容，以保证识别与翻译的准确性。软件由两大核心功能构成：

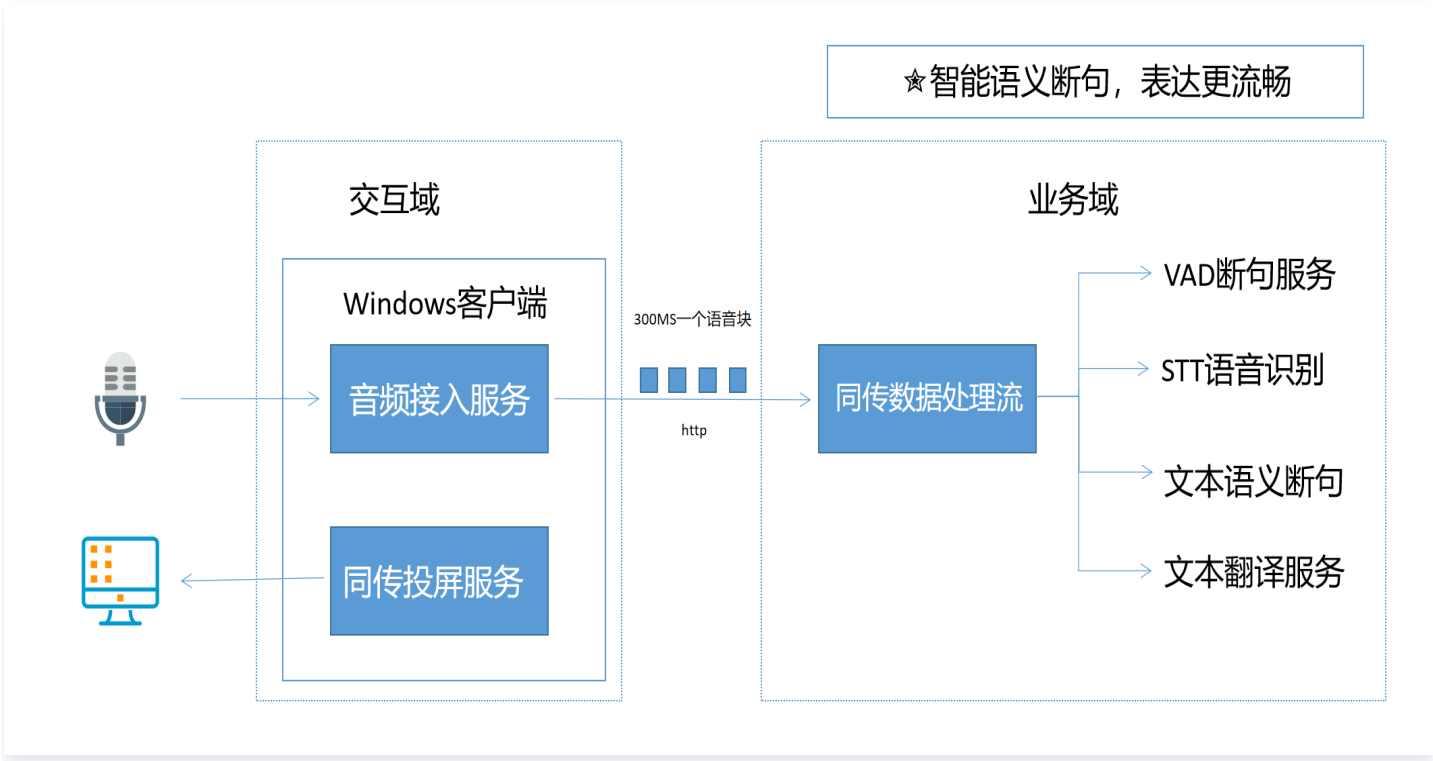
- 基于音频的流式识别，实时获取转写、翻译结果。该功能侧重实时语音输入、识别/翻译结果实时获取、同步展示投屏、同步输出会议纪要等即时性需求。
- 基于音视频文件上传，异步获取转写、翻译结果。该功能侧重基于音视频文件进行转写，以及二次编辑、整理成稿等非即时性需求。

系统架构

硬件系统结构



软件系统结构



产品功能

实时会议

最近更新时间：2024-08-29 15:13:11

实时会议主要用于实时转写与翻译的场景，通过声卡连接场所内的音像设备实现采音，系统实现对采集音频的实时转写与编辑整理。

会前配置

用户根据会议或业务性质，在新建会议同传时，进行相关配置。包括会议信息及高级设置。

会议信息

- 声源语言：支持选择纯中文/纯英文/中英混合/日语/韩语/粤语发言。
- 翻译显示：会议支持选择显示双语/原文/译文，以适应单语会议场景或跨语言会议场景下的不同需求。

新建实时会议

会议名称：

声源语言：

中文 - 英文 ^

翻译显示：

中文 - 英文

原文

显示译文

声源测试：

英文 - 中文

中文 - 日文

检测声源

声源设备：

日文 - 中文

中文 - 韩语

韩语 - 中文

中文 - 粤语

耳机 (ZOMBIESCAT-NC1 Hands-Free A)

角色分离：

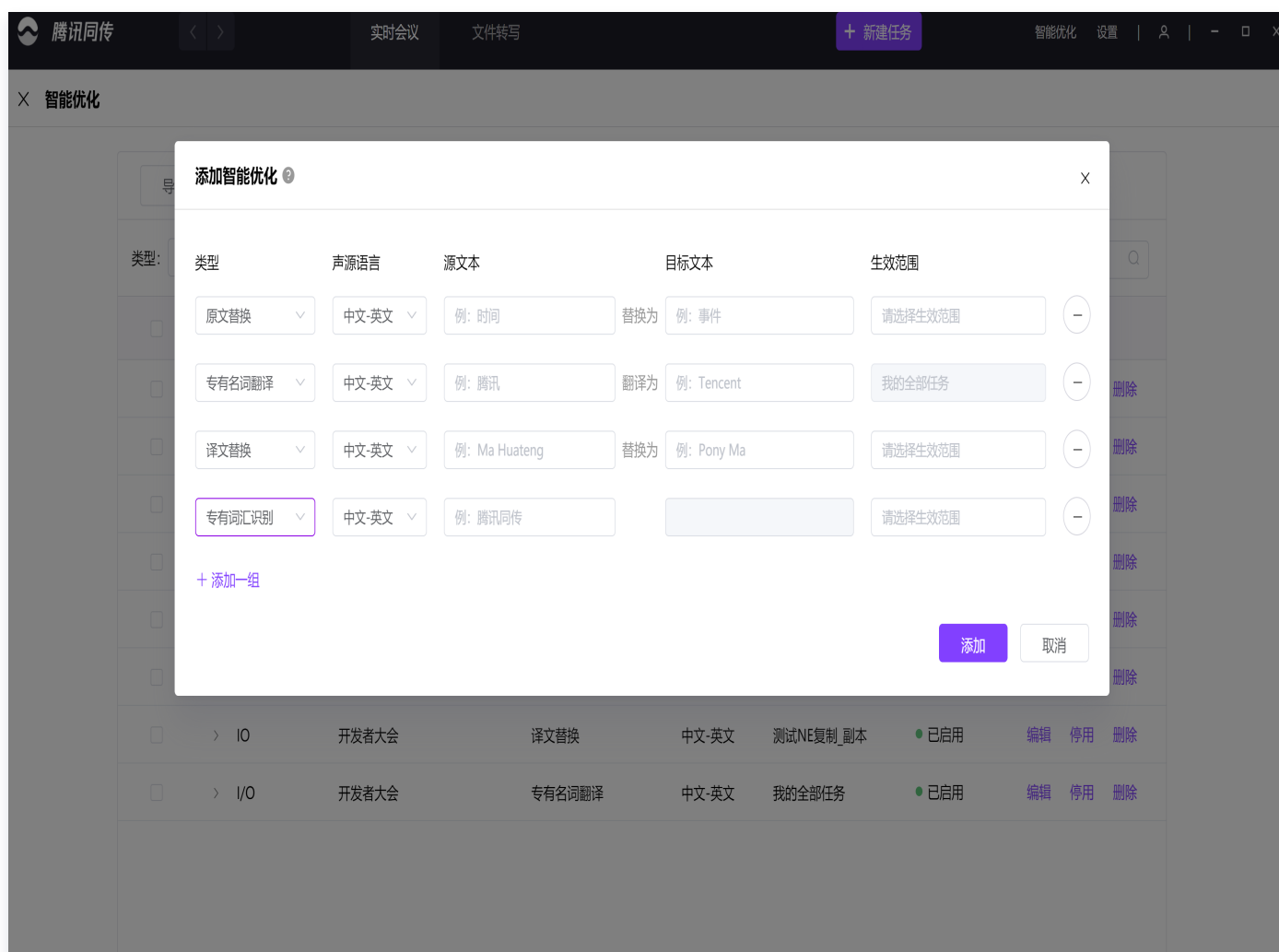
开启角色分离 ?

新建

取消

高级设置

- **字幕投屏**：支持会议选择是否启用投屏功能，对应匹配不同会议视图，以满足不同会议场景下的需求。
- **智能优化**：可针对单场会议定制专属的优化任务，包含原文替换、译文替换、专有名词翻译、专业词汇优化、敏感词过滤等，会前预先定制有效提升识别准确率和翻译准确率。
- **声源选择**：基于 Windows 系统产品支持声源选择，Linux 系统暂不支持。声源选择支持内录、本地麦克风和外接声卡等方式的单一选择和组合选择。
- **内录**：可满足将电脑音视频内容实时转写翻译成文字的需求。
- **本地麦克风**：适用于近场会议或业务沟通场景。
- **外接声卡**：适用于远场会议沟通场景。
- **内录+本地麦克风/外接声卡**：适用于远程在线会议沟通场景。
- **角色分离**：在新建会议时，可根据外接麦克风的顺序配置对应的发言人昵称。会议开始过程中，根据配置将在发言内容前展示发言人昵称，实现角色分离。支持在会前、会中修改和删除发言人。



会中功能

- **会管会控**：用户基于客户端，即可自主控制会议进程（开始/暂停/结束等）。对于多语混合发言的会议，支持用户在会中通过切换发言人语言（说中文/说英语/说日语/说韩语/说粤语），帮助 ASR 引擎判断识别语种，切换到识

别结果实时生效。

- **实时语音转写**：系统通过专业麦克风对全程进行高保真录音，并针对连续语音流，进行实时语音转写识别与翻译，并持续进行转写结果文本内容的输出。
- **实时中英文翻译**：在实时的会议过程中，系统支持实时将发言人的发言结果翻译为目标语种，打破涉外沟通场景下的语言障碍，有效提升涉外会议场景下的沟通效率和质量。
- **实时编辑**：支持对实时转写和翻译的结果进行灵活编辑，包括删除、新增、修改识别/翻译错误的文本、合并/拆分段落等。在对识别文本进行修改时，系统将根据修改结果，自动调取机器翻译引擎，完成翻译结果的同步修改。
- **实时优化**：会中支持增、删、改单场会议的优化任务，包含原文替换、译文替换、专有名词翻译优化、词汇通俗化替换、不断句等。

实时优化

针对本场会议，对后续的认可或翻译结果及时修正

×

类型	源文本	替换为	目标文本
原文替换	例：时间	替换为	例：事件
专有名词翻译	例：腾讯	替换为	例：Tencent
词汇中间不断句	例：更高更快更强	替换为	—

+ 添加一组

取消

确定

- **会议纪要**：系统支持在会议同传过程中，同步输出会议纪要，以帮助节省会后整理文稿的时间、提升开会效率。点击**纪要图标**，屏幕右侧展开类便签样式的会议纪要编辑区，用户可参照左侧实时转写结果，将会议重点内容复制粘贴到右侧，再进行提炼编辑。左侧转写结果区域和右侧会议纪要区域大小支持调节，以满足用户不同场景下的编辑需求。



- **自动摘要：**系统利用自然语言理解技术，从原始结果中提取重点内容，辅助进行会议重点纪要的整理。切换到**自动摘要**中并打开自动摘要，此时进入自动摘要过程，用户可自由控制开启/关闭，会议结束后可对摘要内容进行手动编辑修改及导出。



- **同步投屏：**在会议同传过程中，若需将同传实况通过外接屏实时展示给与会人员，可开启腾讯同传客户端的投屏功能。用户可自由设置投屏语言、投屏样式、字体颜色、字幕背景等，以满足不同场合下多样化的投屏需求，包括支持在进行 PPT 宣讲时提供半屏字幕展示、发言/宣讲时全屏投屏，包括多语言交流场合提供中英双语投屏。

- 协同编辑：在会议中，可添加协同编辑人，协同编辑人可以在会议中进行编辑会议内容。



会后功能

- 会后编辑：支持会后对转写和翻译的结果进行灵活编辑，包括删除、新增、修改识别/翻译错误的文本、合并/拆分段落等。在对识别文本进行修改时，系统将根据修改结果，自动调取机器翻译引擎，完成翻译结果的同步修改。
- 全文搜索替换：单击**全文替换**，在查找文本框输入文字即可对全文搜索，所查找的关键词高亮展示；在替换文本框输入需要替换的文字，单击**替换**可按全文由上到下逐一替换，单击**全部替换**可直接替换全文所有关键词；单击**添加到全局[智能优化]**替换任务自动加入至全局优化，后续所有实时会议及文件转写全部执行此任务。



- 会议内容导出：单击**导出**，系统支持将会议同传、会议纪要、自动摘要、会议音频内容进行导出。会议记录导出格式，支持 word、txt 格式；会议纪要及自动摘要默认为 word；会议音频默认为 wav 格式。

- **音频回放：**为了便于会后编辑，可使用音频回放功能。会议结束后，单击**音频回放**，系统自动将录制的音频按顺序进行播放，此时可进行边听边改操作。也可以将鼠标键入到转写结果的某一处，点开音频回放，系统可从键入文本所在的句子开始向后播放。
- **按句回听：**会议结束后，如有部分语句需要针对性再回听的，可以用鼠标滑到文本上，点击播放按钮即可回听到对应的音频，回听音频所对应文本加粗显示，方便寻找音频对应的文本。



文件转写

最近更新时间：2024-08-23 10:25:11

上传音视频文件

除会议同传外，软件系统还支持将音视频文件快速转写成文字，并同步翻译，对结果进行简单校正即可成稿。

- 支持显示语言（双语/原文/译文）的选择。
- 支持批量上传多个文件。
- 支持 mp3、.mp4、.flv、.mxf、.avi、.mov、.wav、.m4v、.mkv 等常见音视频格式的文件转写。
- 支持查看转写中的文件进度以及完成剩余时间。

上传音视频文件

* 文件语言

☐ 中文

☐ 英文

+

点击上传音视频文件

支持mp3、mp4、flv、mxf、avi、mov、wav、m4v、mkv;
单个音频≤500M，单个视频≤2G

取消

提交

任务插队

为应对突发或紧急任务的插入，系统支持任务插队功能。插队任务可优先进入转写队列。任务插队功能仅开放给管理员账号使用，普通成员不可见。

新建文件转写任务，上传提交后，通过页面左侧缩略列表，单击**更多**进入列表详情页。找到希望使用插队功能的任务，单击**优先转写**即可插队。如需求有变，在进入转写队列前，可通过相同路径单击**取消优先**，即可取消插队。

插队任务无数量限制。仅限于针对“数据准备中”和“转写中”两种状态的任务使用，“正在转写中”的任务不受插队影响。由于算力限制，并行转写任务数有上限，使用插队功能可能对部分任务产生影响，建议慎重使用。

历史文件管理

最近更新时间：2024-08-23 11:39:06

系统将会议同传以加密形式保存在云端，文件转写的结果文件存储在本地，保障内容的保密性，并方便进行统一回溯管理。用户可通过历史文件列表随时查询检索文件。

音视频转写 / 全部文件列表

进行中 4 已完成

全部创建人

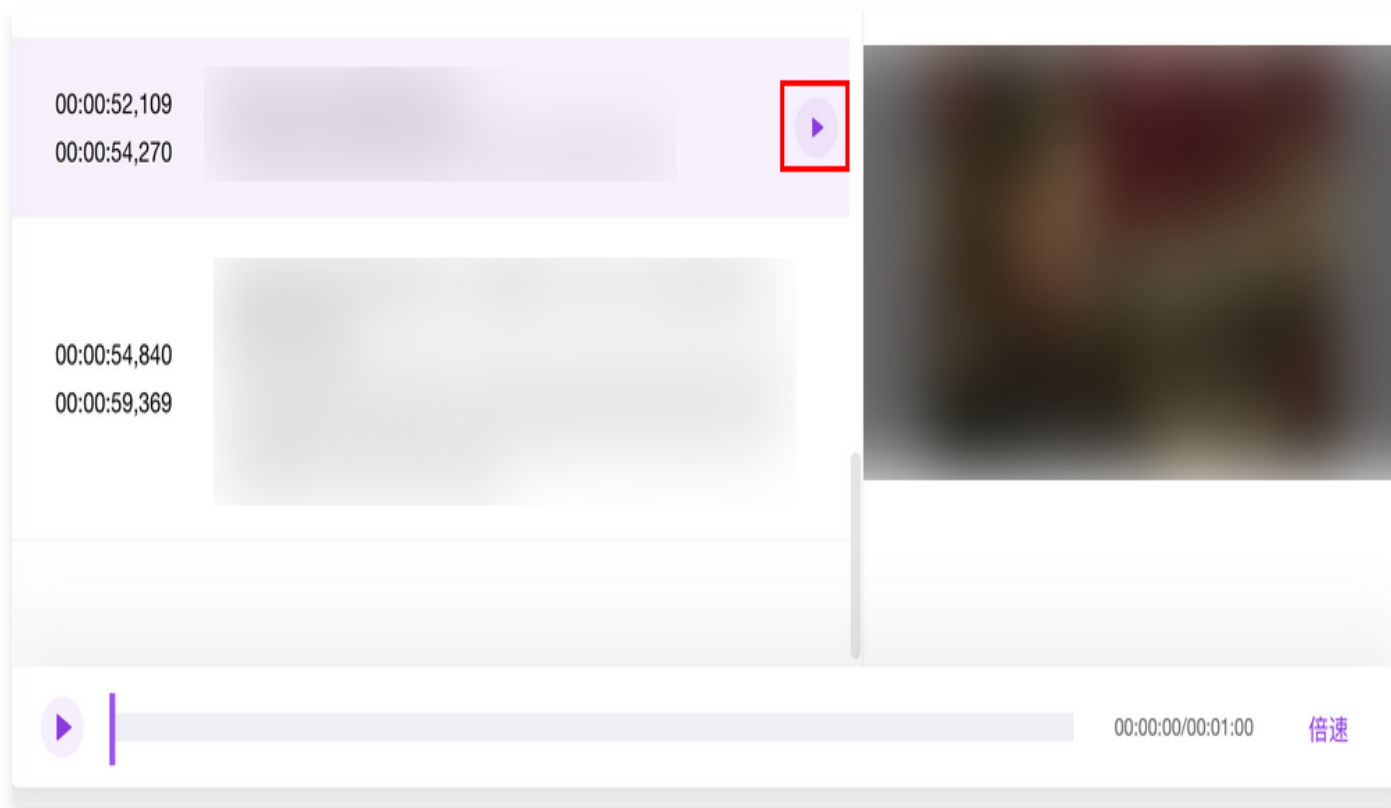
文件名称	语言	创建人	创建时间	状态	操作
	中文		2020-02-27 12:20	数据准备中...	删除
	英文		2020-02-27 12:19	转写中 40% 剩余 00:03:10	删除
	英文		2020-02-27 12:17	转写中 80% 剩余 00:01:57	删除
	英文		2020-02-27 12:14	等待中 1	删除

编辑校正

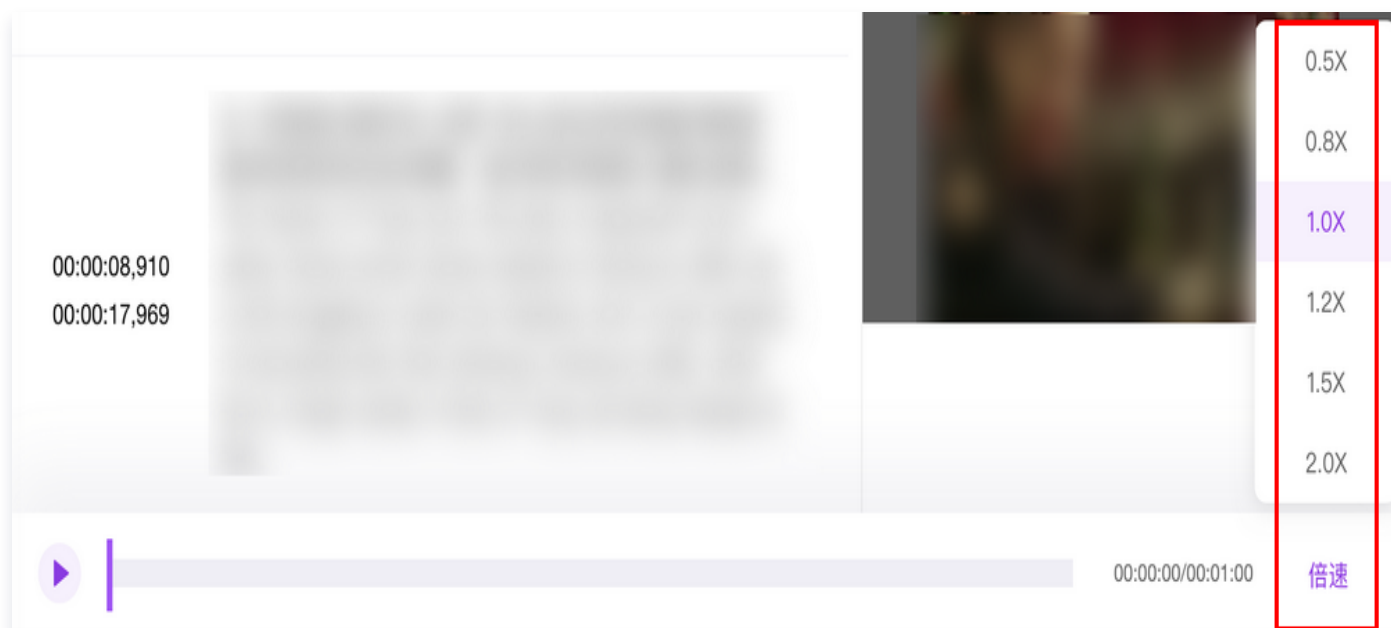
通过历史文件列表，可对会议同传、文件转写文件进行二次编辑校正。

- 为提升编辑效率，系统支持音频的顺序播放和跳播。将鼠标键入文本内容某处，或拖动音频进度条，均可使音视频内容和文本内容同步跳转到相应位置。

- 如有部分语句需要针对性回听时，可以用鼠标悬停在对应文本处，点击右侧播放按钮，即可回听对应的音频。



- 为辅助文档校验并提升校验效率，系统提供音频倍速调节的功能，现已支持0.5x/0.8x/1.0x/1.2x/1.5x/2.0x等6个档位的倍速，用户可根据不同文件下的音频表现自由调节，调节后视频画面和音频同步生效。



文件导出

单击**导出**，系统支持将文件转写内容进行导出。

- 导出格式，支持 srt、word、txt 格式。
- 导出内容，支持导出为双语对照、仅原文、仅译文三种语言。

导出选项

×

文件：

格式：

☒ srt

☐ word

☐ txt

导出内容：

☒ 双语对照

☐ 仅原文

☐ 仅译文

取消

确定导出

智能优化

最近更新时间：2024-08-23 14:45:25

在系统使用过程中，针对通用型的专有名词、网络热词、人名、地名、产品名、缩写简写等，系统提供了优化功能，可根据需要进行预先定制优化，秒级生效，能有效提升识别准确率和翻译准确率。**优化功能包括翻译修正、文本优化、专有词汇识别。**

- **翻译修正：**主要包括原文替换、译文替换、专有名词翻译优化、词汇通俗化替换。
- **原文替换：**将引擎识别不准确的内容替换为正确的识别内容，替换后，翻译结果会根据修改后的识别内容进行更新。
- **译文替换：**将引擎翻译不到位的内容替换为所需的翻译内容，替换后，不会影响相关识别结果。
- **专有名词翻译优化：**适用于业务场景中经常会涉及到的特殊人名、地名、专有名词等的翻译优化。
- **词汇通俗化替换：**若会议内容包含缩略名词、网络热词时，建议选择此修正类型，将其替换为完整、标准的词汇，以保证翻译准确性。
- **文本优化：**包括中间不断句、结尾不断句。
- **中间不断句：**将引擎识别句中中断句错误或特定不能断句的专业内容，修正为不断句。
- **结尾不断句：**将引擎识别结尾断句错误或特定结尾不能断句的专业内容或标点符号，修正为不断句。
- **专有词汇识别：**支持客户自主上传专有词汇，如品牌名、行业特定词汇，从而显著提升识别准确率。

腾讯云

实时会议文件转写

+ 新建任务

智能优化 设置

×

智能优化

搜索源文本/目标文本/创建人

类型：请选择

声源语言：请选择

生效范围：请选择

源文本	目标文本	类型	声源语言	生效范围	操作
> 为准		原文替换	中文	所有人的全部任务	编辑 删除
> 厉害	威力	专有名词翻译	中文	所有人的全部任务	编辑 删除
> 结束		词汇末尾不断句	中文	所有人的全部任务	编辑 删除
> 形成风		词汇中间不断句	中文	所有人的全部任务	编辑 删除
> Tencent	tencent	译文替换	中文	所有人的全部任务	编辑 删除
> 开始	现在	原文替换	中文	所有人的全部任务	编辑 删除

产品优势

最近更新时间：2024-08-23 10:56:01

准确率高

腾讯同传引擎由腾讯核心 AI 团队（微信智聆与腾讯翻译君）联合打造。

微信智聆提供语音识别功能，将音频转成文字，通过先进的区分性训练方法进行语音建模，使语音识别引擎在复杂应用环境下均有良好的效果表现，其中普通话识别准确率可达97%。

腾讯翻译君提供机器翻译功能，其核心引擎 NMT 基于目前先进的 Transformer 模型，并融合新型循环神经网络，在提升了训练效果的同时，翻译质量更佳，翻译速度更快。此功能在 2018WMT 国际翻译大赛中获得中英方向世界第一，翻译可接受度超92%。

安全稳定

腾讯同传服务久经业务考验，微信智聆的语音识别服务，日均请求量4亿次，日均处理行业语音5万小时。腾讯翻译君提供的翻译服务，日均翻译请求5亿次，方案部署各垂直行业，在复杂应用环境下均有良好的识别与翻译效果，产品的技术稳定性与方案成熟度得到了市场验证。

自定义优化

腾讯同传提供可视化训练配置页面，用户根据其业务场景自定义完成语音识别、机器翻译结果的优化，满足不同行业定制化的语言需求，有效提升输出结果的准确率。

应用场景

最近更新时间：2024-08-23 11:28:11

会议实时记录

在企业、政府的日常工作会议中，根据三重一大议事原则，会议必须全程留痕且需要归档备查，当前多采用人工记录的方式，存在出稿效率低、内容不完整、纸质文件后期引用难等问题。通过腾讯同传可以实时记录、实时编辑会议记录，可以直接导出 word 文档方便会议存档。

涉外会议实时翻译

在涉外会议中，会议双方由于语言障碍，通常需要外聘翻译人员进行翻译，在增加了会议费用成本的同时，又扩大了会议内容的传播范围。腾讯同传利用先进的机器翻译技术，可以实时进行中英互译，有效降低了中外双方的沟通难度，减少会议的成本和控制了会议内容的传播范围。

会议字幕实时上屏

在大型会议下，现场存在嘈杂、声音小、讲话人语速过快等因素，影响观众有效获取信息，尤其是涉及双语交流的场景。通过腾讯同传的实时上屏功能，只需一台电脑就可直观展示宣讲内容与翻译内容，减少会议方对同声传译、同传设备的费用开支。

个人速记

企业领导、政府官员在日常交流和会议中，经常会口述信息，需要秘书对内容进行文字记录；作家、编剧、记者等文字职业，也会存在大量的需要将口述内容变成文字的诉求。这两类人群，当前都是通过对口述内容进行录音，事后再回听录音并手动生成文字，通过腾讯同传可实时记录口述内容，也可以会后对录音文件转写直接生成文稿，提高个人工作效率。

录音转写

针对通过其他录播系统或录音设备获取的非实时的音频文件，系统支持导入历史录音进行快速转写，1个小时的音频文件约6 - 8分钟即可完成转写，且转写结果会自动分段，支持转写结果按句播放，提升转写结果的可阅读性和编辑效率。

核心技术

最近更新时间：2024-08-23 17:24:07

语音识别引擎

语音识别技术（Automatic Speech Recognition）是一种实现从“声音”到“文字”转换的技术，通过将人的语音直接转换成相应的文本以便计算机进行理解和产生相应的操作，并最终实现人与机器之间的自然语音交互。

腾讯同传采用国内广泛应用的微信智聆语音识别引擎，针对语音识别应用中面临的方言口音、背景噪声等问题，基于开放业务系统中所收集的涵盖不同方言和不同类型背景噪声的海量语音数据，通过先进的区分性训练方法进行语音建模，使语音识别引擎在复杂应用环境下均有良好的效果表现。

腾讯同传所应用的引擎具备的前端特性如下：

- 支持中/英文语音转写

语音识别对于日常使用的常用对话有着很高的识别准确率，包含短信类、生活、交通、娱乐、科技、数字数值、名人、互联网热词、新闻等领域，其中中文标准普通话语音识别准确率最高可达97%以上，Native 英文整句识别准确率可达到93%以上。

- 支持标点智能预测

语音识别使用超大规模的语言模型，对识别结果语句智能预测其对话语境，提供智能断句和标点符号的预测。

- 端点检测

端点检测是对输入的音频流进行分析，确定用户说话的起始和终止的处理过程。一旦检测到用户开始说话，语音开始流向识别引擎，直到检测到用户说话结束。这种方式使识别引擎在用户说话的同时即开始进行识别处理。

- 噪音消除

在实际应用中，背景噪声对于语音识别应用是一个现实的挑战，即便说话人处于安静的办公室环境，在语音中也难以避免会有一些的噪声。语音识别系统应具备高效的噪音消除能力，以适应用户在千差万别的环境中应用的要求。

- 大词汇量、独立于说话人的识别功能

满足大词汇量、与说话人无关的识别要求。产品可以支持数万条语法规模的词汇量，并能适应不同年龄、不同地域、不同人群、不同信道、不同终端和不同噪声环境的应用环境。

- 敏感词过滤

接入腾讯统一的敏感词过滤服务，以实现可靠的敏感词过滤，避免实时会议展示敏感内容。

本次系统具备的前端特性如下：

- 词汇识别系统

满足大词汇量、与说话人无关的识别要求。可以支持数万条语法规模的词汇量，并能适应不同年龄、不同地域、不同人群、不同信道、不同终端和不同噪声环境的应用环境。

- 置信度输出

置信度反映了识别结果的可信程度。语音识别引擎可以在返回识别结果时携带此识别结果的置信度，应用程序可以通过置信度的值进行分析和后续处理。

- 多识别结果

又称多候选技术，在某些识别过程中，识别引擎可以通过置信度判决的结果向应用程序返回满足条件的多个识别

结果，而不是唯一的结果。识别系统提供了可能的识别结果列表，并按置信度结果从高到低进行排列。在业务设计中，可以根据应用要求向工作人员提供这些结果，供工作人员进行二次选择。通过置信度判决和多识别结果输出技术可以开发更加灵活、更加人性化的业务流程。

- 热词识别系统

热词识别使得语音识别应用程序能够在说话者说话的同时检测一个特定的词或短语。

机器翻译引擎

核心引擎 NMT 基于目前先进的 transformer 模型，并融合新型循环神经网络，在提升了训练效果的同时，翻译质量更佳，翻译速度更快。有效结合序列自注意力模型和循环网络模型的优势，实现了两个模型的优势互补，在保证翻译质量的前提下，提升翻译速度，能够满足更多的实时性翻译需求。

应用引擎翻译能力包括如下几种：

- 支持进行本地上传所需定制的词汇内容，针对敏感度高、优先级高的重要特殊词汇（例如领导人名字、政府机构、地名、缩写简称、专有词汇、行业术语等）秒级生效。
- 具备对文稿中专有名词、缩写简称、行业术语等特殊词汇的处理能力，并正确翻译成指定内容。例如一带一路、APEC、两会等。
- 针对语言风格、翻译记忆等需要长期时间与大量数据积累的内容，可提供上传与保存工具，积累到一定量级后，由专业人员进行现场训练。
- 可处理混合在一起的文稿，并正确翻译成指定语种。
- 可自定义关键词、关键字，并设定翻译结果。可自动识别文本中自定义的关键词、关键字，并按指定结果进行翻译。