

AI_DH 文档



【版权声明】

版权所有©百度在线网络技术（北京）有限公司、北京百度网讯科技有限公司。 未经本公司书面许可，任何单位和个人不得擅自摘抄、复制、传播本文档内容，否则本公司有权依法追究法律责任。

【商标声明】



和其他百度系商标，均为百度在线网络技术（北京）有限公司、北京百度网讯科技有限公司的商标。 本文档涉及的第三方商标，依法由相关权利人所有。未经商标权利人书面许可，不得擅自对其商标进行使用、复制、修改、传播等行为。

【免责声明】

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导。 如您购买本文档介绍的产品、服务，您的权利与义务将依据百度智能云产品服务合同条款予以具体约定。本文档内容不作任何明示或暗示的保证。

目录	
目录	2
产品简介	4
产品概述	4
应用场景	4
平台操作指引	13
Chrome麦克风权限教学	16
购买指南	18
价格指南	18
退费说明	21
数字人形象使用	22
形象类别介绍	22
2D小样本数字人	24
3D数字人	89
2D精品数字人	105
端渲染数字人	108
公共音色库	111
组件通用服务	119
接口通用说明	119
SSML使用说明	125
文件上传接口	125
定制服务	128
照片数字人形象包定制接口	128
2D小样本数字人（通用口型）拍摄指南	134
2D小样本数字人（专属口型）拍摄指南	139
音色克隆和查询接口	144
2D小样本数字人形象定制-通用口型接口	150
2D小样本数字人形象定制-专属口型接口	156
端渲染2D小样本数字人拍摄指南	161
2D小样本数字人抠像接口	166
声音克隆LITE版	170
生成式数字人形象定制接口	174
视频组件接入手册	179
123数字人视频	179
123数字人视频合成接口	179
123数字人拍摄指南	182
照片数字人视频合成接口	184
基础视频合成接口	191
高级视频合成接口	197
语音合成接口	209

Baidu 百度智能云文档	产品简介
交互组件接入手册	212
云渲染交互	212
H5版	212
小程序版	231
交互会话管理接口	239
端渲染交互	241
照片数字人交互SDK-H5版	241
2D数字人交互SDK-Android版	247
3D数字人交互SDK-H5版	274
3D数字人交互SDK-Android版	282
3D数字人交互SDK-iOS版	302
手语数字人交互SDK-Android版	310
云渲染直播	321
云渲染推流接口	321
联系我们	329
曦灵数字人定制组件克隆协议	329
MCP服务	331
数字人MCP服务使用说明	331

产品简介

产品概述

曦灵数字人开放平台面向企业及个人开发者，提供高性能、易集成和多样化的数字人组件能力，满足数字人形象定制、视频合成、交互对话、直播等应用需要，支持公有云调用与私有化部署。帮助开发者快速集成数字人能力，助力企业打造全场景应用。智能集成，轻松搞定。

🔗 产品优势

- **高性能**
人像训练、视频合成和单图驱动效率领跑行业
- **易集成**
轻量级SDK和标准化接口助力快速集成
- **多样化**
多渠道交互适配，多渲染方案提效，多功能组件开放

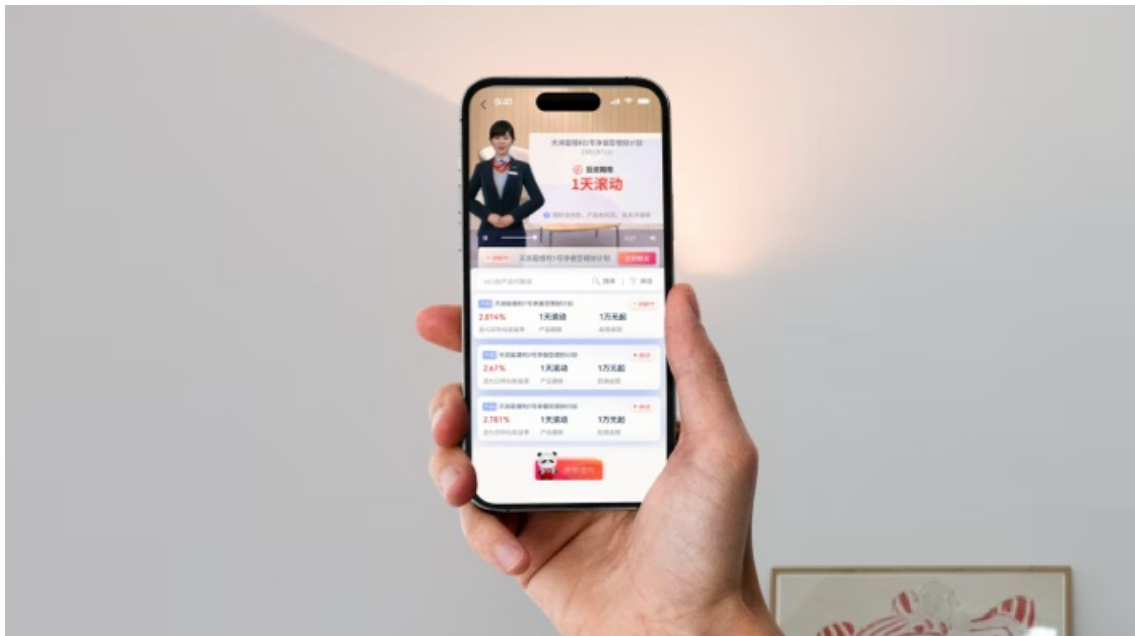
🔗 产品能力

- **数字人对话**
 - 云渲染数字人：快速搭建数字人对话应用，支持2d精品数字人、3d数字人等数十款形象免费用
 - 端渲染数字人：快速集成数字人终端渲染，全场景覆盖多终端适配
 - 照片数字人：快速打造千人千面的对话体验
- **数字人播报**
 - 云渲染视频：输出2D/3D数字人高质量视频
 - 小样本数字人：极速制作数字分身，便捷调用集成
 - 照片数字人：快速生产千人千面的数字人视频
- **数字人直播**
 - 直播推流API：数字人直播在线实时渲染能力，效果还原真人，人像风格众多、场景真实，能够满足大多数的直播场景需求
- **通用能力**
 - 声音克隆：高质量还原声音本色，打造音容如真的数字人
 - 人像定制：支持3D高精超写实数字人、真人拍摄定制服务

应用场景

🔗 金融

- **数字理财专员**
通过数字人进行理财产品介绍，在线解答理财产品的相关问题，提供购买引导和咨询，给予用户最佳理财建议
相关组件：云渲染数字人、端渲染数字人、云渲染视频、单图交互
案例：



- 远程坐席

与数字人面对面沟通，提供远程银行业务咨询与办理指导，帮助用户自主完成业务办理，降低服务人力成本。

相关组件：云渲染数字人

案例：



- 智能客服

建设7×24小时服务的数字人营业厅，为用户提供全时段的智能自助服务，全方位解答用户的疑问。

相关组件：云渲染数字人、端渲染数字人

案例：

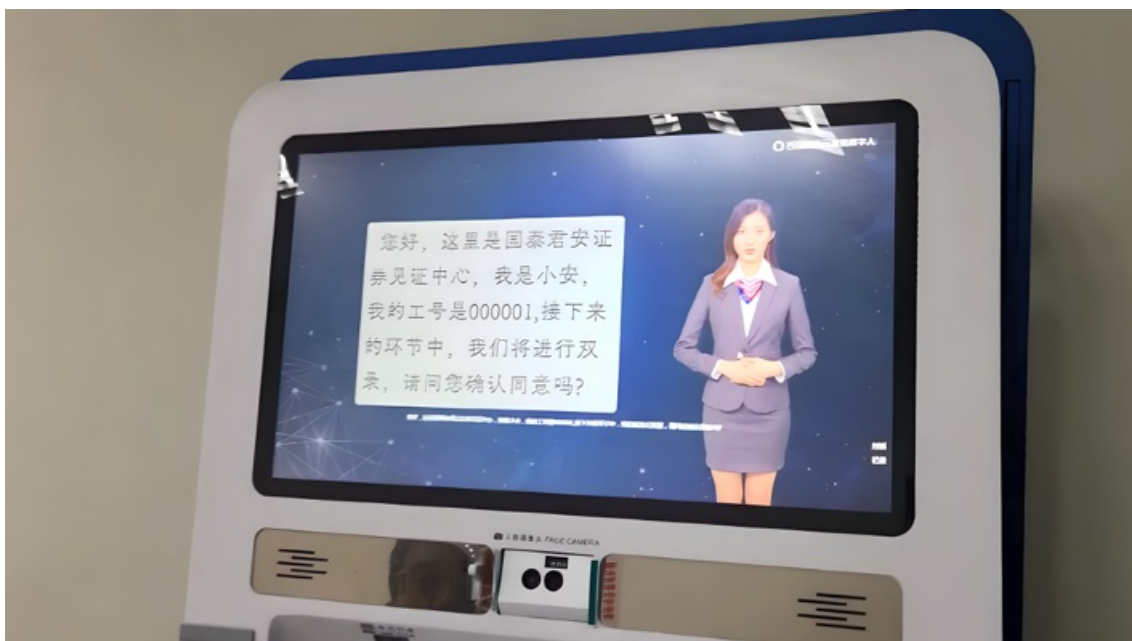


- **VTM智能柜员**

打造数字人智能柜员，为用户提供最有效的业务办理引导和风险提醒，帮助用户便捷办理业务。

相关组件：端渲染数字人

案例：



- **网点大屏一体机**

提供比拟真人的数字人交互体验，在网点业务办理过程中实现前置分流，为用户指引最佳路径。

相关组件：端渲染数字人

案例：



- 多功能视频制作

为营销、宣传和培训等多场景提供数字人视频能力，帮助打造低成本高效率视频制作体验。

相关组件：云渲染视频、分身视频、单图驱动视频、声音克隆

案例：



政务

- 政务服务助手

数字人+政务知识为用户提供准确的咨询服务，解决政务服务最后一公里难题，全天候陪伴式服务。

相关组件：云渲染数字人、端渲染数字人

案例：



● IOC智能问政

基于数字人和大模型能力打造智能问政能力，实现IOC大屏的操控、调度、分析和问答等多种能力，提升IOC智能化水平。

相关组件：端渲染数字人

案例：



● 展厅数字人

建设智能化展厅问答数字人，实现展厅讲解自助化和智能化，有效降低讲解人力成本。

相关组件：云渲染数字人、端渲染数字人

案例：



- 政务办事大厅大屏一体机

提供比拟真人的数字人交互体验，在政务业务办理过程中实现问题速查与智能问答，详细精准的为用户答疑解惑。

相关组件：端渲染数字人

案例：



- 多功能视频制作

为新闻、宣传、反诈和培训等多场景提供数字人视频能力，帮助打造低成本高效率视频制作体验。

相关组件：云渲染视频、分身视频、单图驱动视频、声音克隆

案例：



医疗

• 数字医生

基于数字人能力打造具备在线诊疗能力的数字医生分身，全天候为患者提供专业咨询支持

相关组件：单图交互、分身交互、声音克隆

案例：



• AI药品说明书

提供可视化可交互的数字人AI药品说明书，AI技术加持适老化改造，帮助中老年患者轻松获取药品

相关组件：云渲染数字人、端渲染数字人、云渲染视频

案例：



- 健康科普
基于医生数字分身快速生产健康科普视频，提升视频制作效率和人力时间占用。
相关组件：分身交互、声音克隆
案例：



教育

- 口语对练
快速搭建口语对练数字人场景，与数字人对练，为学生提供全时在线的专属对练体验。
相关组件：云渲染数字人、端渲染数字人
案例：



- 授课视频

打造高还原度的教师数字分身，低成本快速制作数字人授课视频。

相关组件：分身视频、声音克隆

案例：



- 在线问答/课后辅导

在线实时帮助学生解答日常学习问题，通过输入专业授课知识，灵活根据学生提出的问题作出解答。

相关组件：云渲染数字人、端渲染数字人

案例：



平台操作指引

百度智能云**曦灵数字人**开放平台面向企业及个人开发者，提供高性能、易集成和多样化的数字人组件能力，满足数字人形象定制、视频合成、交互对话、直播等应用需要，支持公有云调用与私有化部署。帮助开发者快速集成数字人能力，助力企业打造全场景应用。

地址：<https://xiling.cloud.baidu.com/open/overview>

账号：使用已完成认证的百度智能云账号密码登录

激活：登录平台后可以领取权益体验组件功能试用，免费额度消耗完后需在「组件商店」进行购买续费

🔗 平台功能主要包含四大模块：

- 概览
- 组件商店
- 组件管理
- 应用管理
- 数据统计

🔗 平台人像分为：

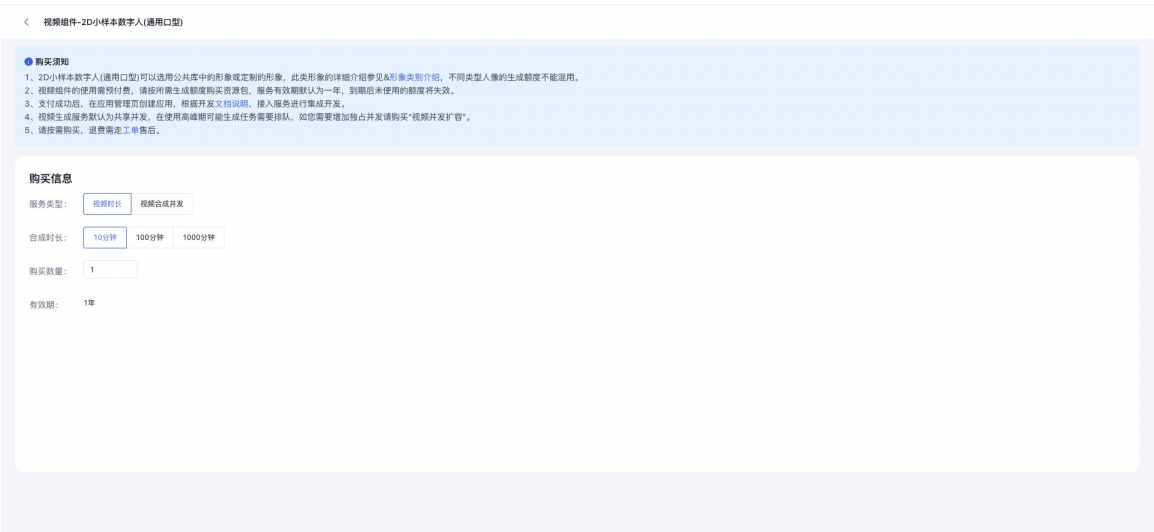
- 照片数字人
- 2D小样本数字人
- 2D精品数字人
- 3D数字人

🔗 一、组件商店选购

用户登录平台后，可以进入「组件商店」了解各项组件能力并查看相关文档

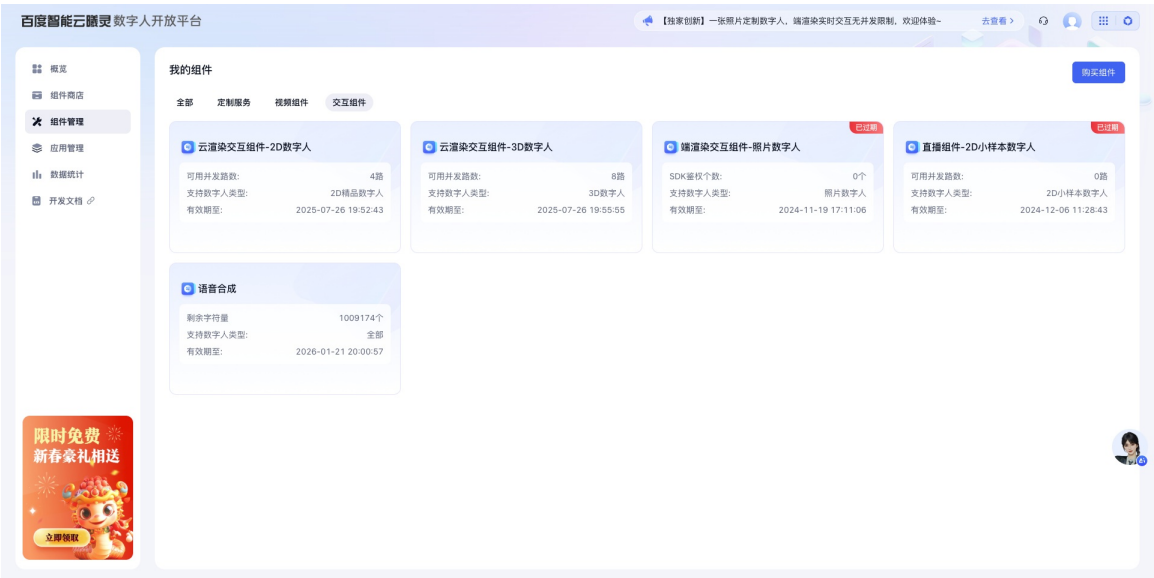


在购买页中，用户可根据需要选择组件类型和对应数量

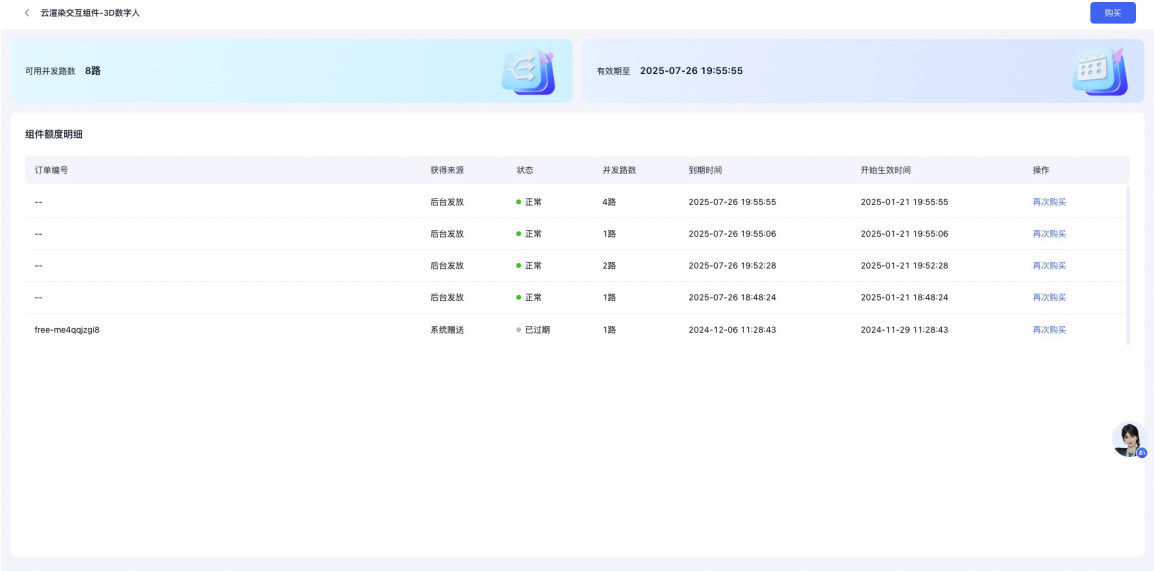


二、组件管理

用户完成组件商品下单支付后，可在「组件管理」中查看已购买组件以及剩余时长、有效期等信息

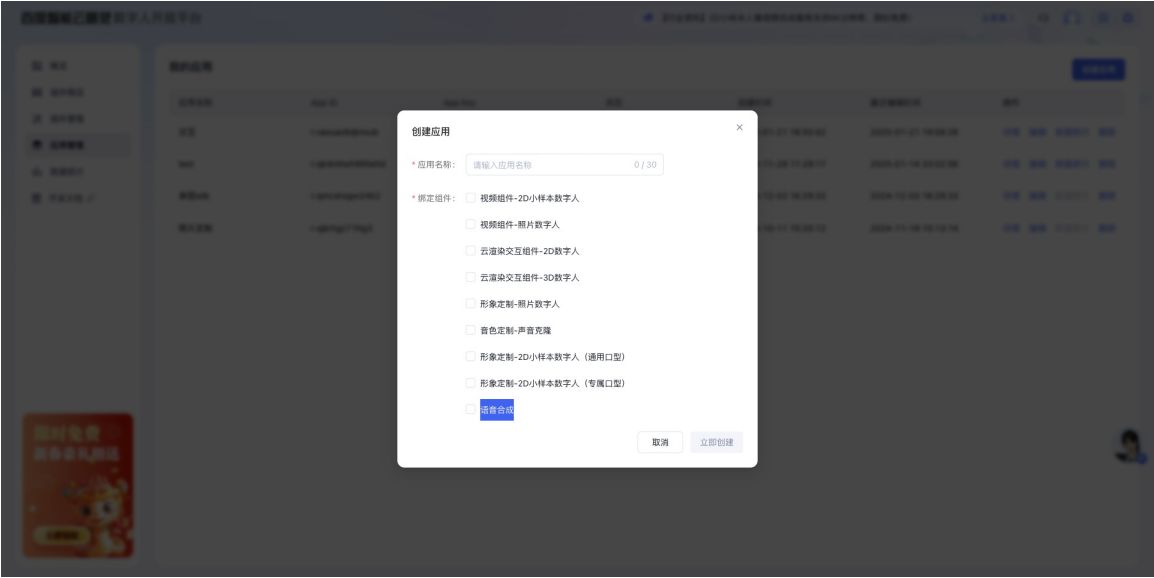


点击“查看详情”可查看订单的具体信息

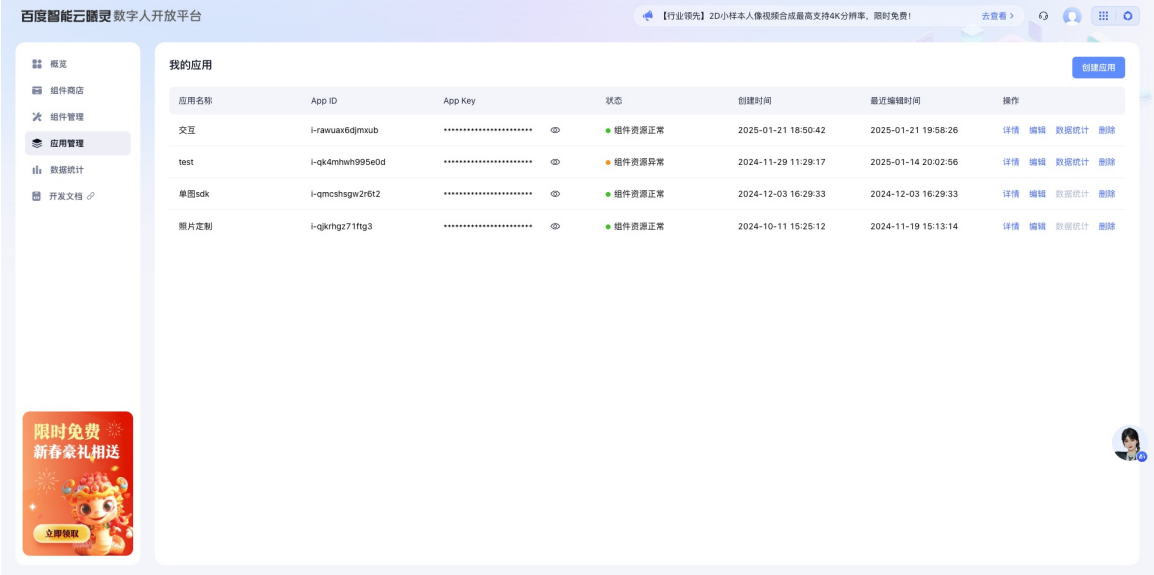


三、应用管理

用户完成组件商品购买后，可进入「应用管理」进行配置，编辑应用名称并绑定所需组件能力

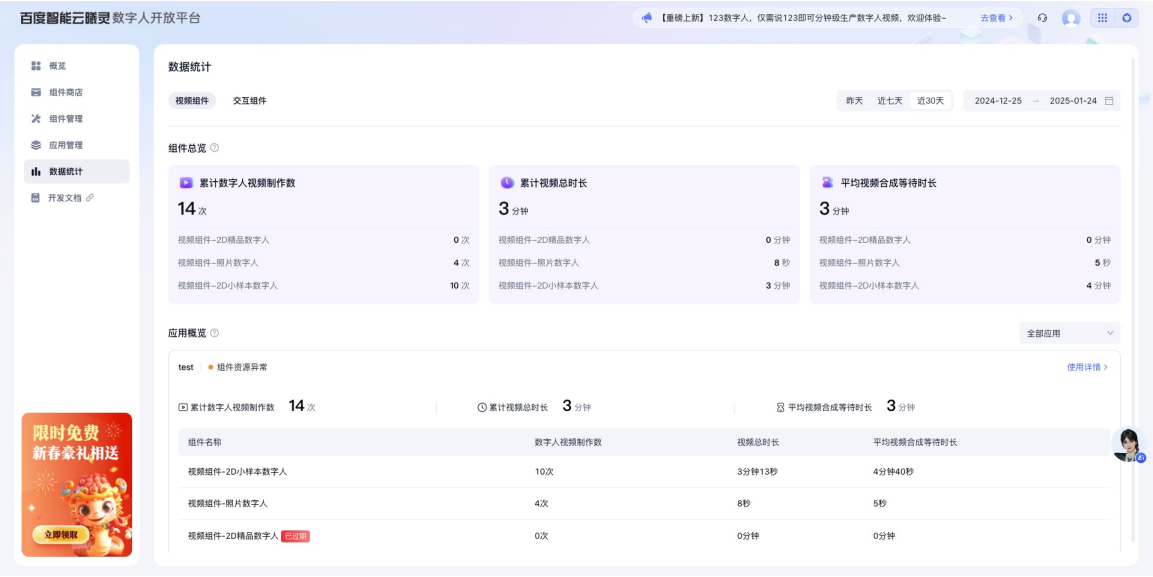


完成创建后可在页面中获取应用鉴权码，通过专属鉴权码进行数字人能力调用

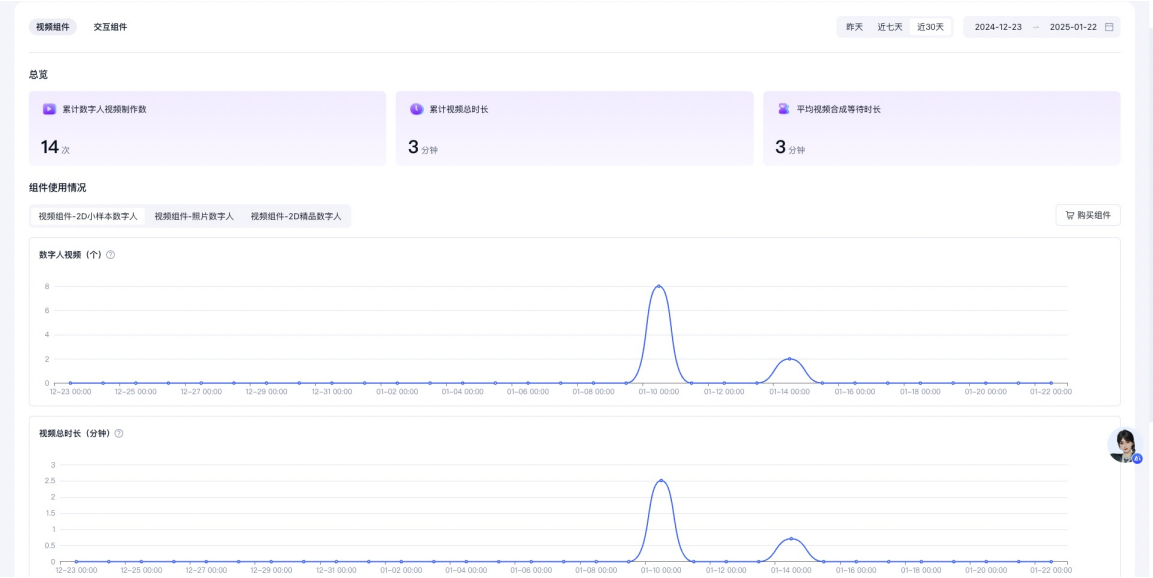


四、数据统计

用户可以在数据统计中查看组件调用量、服务时长等相关信息



完成创建后可在页面中获取应用鉴权码，通过专属鉴权码进行数字人能力调用



Chrome麦克风权限教学

一、前言

本教学文档旨在指导用户如何在Google Chrome浏览器中重新开启麦克风/摄像头权限。如果您在使用某个网页应用时遇到麦克风/摄像头不可用的情况，可能是因为之前拒绝了权限请求。按照以下步骤，您可以重新授予浏览器访问麦克风/摄像头的权限。

二、重新授予浏览器访问麦克风/摄像头的权限步骤

第一步：Chrome浏览器，点击右上角的三个点图标，选择“设置”。



第二步：在设置页面左侧菜单中，点击“隐私和安全”，选择“网站设置”。



第三步：在“权限”部分，点击“麦克风”或“摄像头”（根据您所需的权限选择打开）。



在麦克风或者摄像头设置页面，您会看到一个网站列表，找到您遇到问题的网站。



将该网站的麦克风/摄像头权限设置为“允许”。



返回该网站并刷新页面，麦克风权限应该已生效。

通过上述步骤，您应该能够成功重新授权Chrome浏览器的麦克风/摄像头权限。如有任何问题，欢迎联系小助手支持。

购买指南

价格指南

单独购买形象采购、播报服务以及交互服务任一产品均无法直接应用于最终的应用场景，需要组合购买。

1 定制服务

服务类型	服务概述	价格
照片数字人定制	提供将一张单帧照片秒级生成2D数字人并用于端渲染照片数字人交互SDK的定制服务。	1次服务包 20元 10次服务包 200元 100次服务包 2000元
2D小样本数字人定制（通用口型）	提供一段1-4分钟自然闭嘴的真人视频，支持将其训练为固定背景的数字人或者透明背景数字人，完成训练的人像支持用于视频制作和直播场景。	1次服务包 1000元 10次服务包 10000元
2D小样本数字人定制（专属口型）	提供一段4-8分钟的真人口播视频，支持将其训练为固定背景的数字人或者透明背景数字人，完成训练的人像支持用于视频制作、对话和直播场景。	1次服务包 2000元 10次服务包 20000元
非凡人像定制	提供客户专属非凡人像数字人定制服务（2D），需要客户提供模特，服务内容包括影视级录影棚拍摄、影视级场务指导、非凡人像数字人建模和效果优化等工作，包括1套服装。	1个人像 300000元
2D精品数字人定制	提供客户专属2D精品数字人定制服务，需要客户提供模特，服务内容包括专业录影棚拍摄、2D精品数字人建模和效果优化等工作，包括1套服装，10个动作。	1个人像 200000元
3D卡通数字人定制	提供客户专属3D卡通数字人定制服务，提供按需制作的3D卡通数字人模型1个，包括1套服装配饰制作、1套发型制作、1套基础动作包和技术美术调试工作。	1个人像 200000元
3D数字人定制	提供客户专属3D数字人定制服务，提供按需控制的客户版权3D数字人人脸1个，提供1套服装配饰制作、1套发型制作、1套基础动作包和技术美术调试工作。	1个人像 600000元

2 视频组件

服务类型	服务概述	价格
照片数字人-生成时长包	提供有效期为1年的照片数字人视频生成时长服务包，默认提供公有资源池，不提供额外单独并发。	10分钟服务包 45元 100分钟服务包 450元 1000分钟服务包 4500元
2D小样本数字人-生成时长包	提供有效期为1年的2D小样本数字人视频生成时长服务包，默认提供公有资源池，不提供额外单独并发。	10分钟服务包 30元 100分钟服务包 300元 1000分钟服务包 3000元
2D精品数字人-生成时长包	提供有效期为1年的2D精品数字人视频生成时长服务包，默认提供公有资源池，不提供额外单独并发。	10分钟服务包 40元 100分钟服务包 400元 1000分钟服务包 4000元
3D数字人-生成时长包	提供有效期为1年的3D数字人视频生成时长服务包，默认提供公有资源池，不提供额外单独并发。	10分钟服务包 50元 100分钟服务包 500元 1000分钟服务包 5000元
123数字人-生成时长包	提供有效期为1年的123数字人视频生成时长服务包，默认提供公有资源池，不提供额外单独并发。	10分钟服务包 30元 100分钟服务包 300元 1000分钟服务包 3000元
照片数字人-并发	提供照片数字人的视频合成专属并发1路	专属并发1个月 4500元 专属并发1年 45000元
2D小样本数字人-并发	提供2D小样本数字人的视频合成专属并发1路	专属并发1个月 2000元 专属并发1年 20000元
2D精品数字人-并发	提供2D精品数字人的视频合成专属并发1路	专属并发1个月 2000元 专属并发1年 20000元
3D数字人-并发	提供3D数字人的视频合成专属并发1路	专属并发1个月 2000元 专属并发1年 20000元
123数字人-并发	提供123数字人的视频合成专属并发1路	专属并发1个月 2000元 专属并发1年 20000元

注：视频计费方式为按服务包计费或按服务包+独占并发计费

3 对话组件

服务类型	服务概述	价格
云渲染交互-2D数字人-并发	提供云渲染交互组件-2D数字人的专属并发1路， 用户可限时免费使用百度标准人像库形象。 该组件同时支持用户使用定制的2D小样本数字人或2D精品数字人。	专属并发7天 800元 专属并发1个月 2400元 专属并发1年 24000元
云渲染交互-3D数字人-并发	提供云渲染交互组件-3D数字人的专属并发1路， 用户可限时免费使用百度标准人像库形象。 该组件同时支持用户使用定制的3D数字人	专属并发7天 1200元 专属并发1个月 3600元 专属并发1年 36000元
端渲染交互-照片数字人-SDK license	提供端渲染交互组件-照片数字人的SDK授权license1个， SDK支持H5/web端使用，授权有效期1年， 用户可选择搭配字符包和TTS并发使用。	授权license 1个 100000元
端渲染交互-2D数字人-SDK license	提供端渲染交互组件-2D数字人的SDK授权license1个， 用户可限时免费使用百度标准人像库形象。 SDK支持安卓APP使用，授权有效期1年， 用户可选择搭配字符包和TTS并发使用。	授权license 1个 3600元
端渲染交互-3D数字人-并发	提供端渲染交互组件-3D数字人的SDK专属并发1路， 用户可限时免费使用百度标准人像库形象。 SDK若需文字驱动，需额外购买语音合成-字符包使用。	专属并发7天 120元 专属并发1个月 360元 专属并发1年 3600元
语音合成-字符包	需搭配端渲染交互组件使用。 同时提供TTS字符调用API，用户可以通过鉴权token调用该API， 从而提供个性化数字人音频合成能力调用，字符包有效期1年。	字符包 10w 1个 100元 字符包 100w 1个 1000元
语音合成-并发扩容	需搭配端渲染交互组件使用。 提供TTS字符调用并发扩容，用户可以通过鉴权token调用该API 并发扩容权限，从而扩充个性化数字人音频合成能力支持的并发数量。	并发扩容 1个月 40元 并发扩容 1年 400元

4 直播组件

服务类型	服务概述	价格
直播API-2D小样本数字人	提供直播组件-2D小样本数字人的专属并发1路， 用户可限时免费使用百度标准人像库形象。 该组件同时支持用户使用定制的2D小样本数字人。	专属并发7天 700元 专属并发1个月 2000元 专属并发1年 20000元

5 声音定制组件

服务类型	服务概述	价格
声音克隆极速版	提供声音定制组件-声音克隆极速版训练机会1次， 完成训练的TTS支持用于数字人视频制作、交互和直播组件。	1次服务包1000元 10次服务包10000元

退费说明

1 退费说明

曦灵数字人开放平台提供两类服务：形象定制、视频组件、交互组件。

- 「形象定制」、「视频组件-视频合成时长」不支持退费。一旦正式采购，不支持退费
- 「交互组件」、「视频组件-视频合成并发」属于包年包月类计费项，支持退费。

您可以按照退费流程进行退款申请，退订金额为订单实付金额，包括现金、赠送金以及现金券的部分支持全额退还，代金券/折扣券不支持退还。

🔗 2 退费流程

百度智能云拥有无理由退款特权。在退款有效期内，用户可以通过提交工单的方式进行退款申请，为了保证退款快速受理，请务必在工单中说明需要退款的订单id。如需退款，请点击[申请退款](#)

数字人形象使用

形象类别介绍

🔗 照片数字人

将一张单帧照片秒级生成2D数字人。适用于需要大批量生成数字人，缺乏底板视频拍摄条件，通过照片克隆人像进行简单内容播报。照片数字人支持用于视频制作、交互对话场景。



🔗 2D小样本数字人

将一段带人像的视频分钟级训练生成2D数字人。训练完的形象口型特征为算法匹配结果，适用于需要真人形象进行内容生产，但无需固定人像，追求人像克隆效率的应用场景。2D小样本数字人支持用于视频制作、交互对话和直播场景。



2D精品数字人

基于摄影棚的绿幕拍摄视频训练的2D真人数字人，通过算法模型和渲染引擎的方式驱动，适用于需要高质量2D人像效果和数字人动作编辑的场景。2D精品数字人支持用于视频制作场景和交互对话场景。



3D数字人

基于原画设计和建模制作的3D数字人，通过算法模型和渲染引擎的方式驱动，适用于需要展示数字人表现力和形象高频变化的场景。3D超写实数字人支持用于视频制作场景、交互对话场景。



2D小样本数字人

2025年6月起新增人像将仅在配置台更新，数字人figureid请从 <https://xiling.cloud.baidu.com/open/config> 获取。

🔗 形象 - figure id

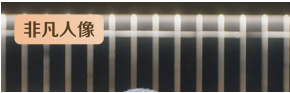
形象适用范围：云渲染交互组件、视频组件、直播组件
如在使用中需调整人像位置，可以在调用组件时通过参数配置人像位移及缩放

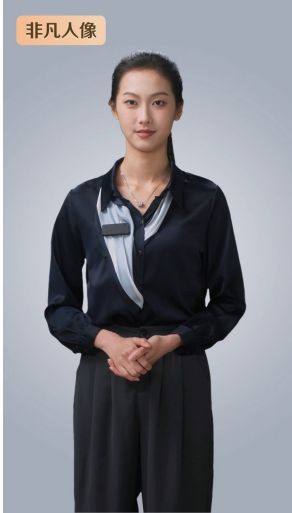
🔗 对话场景

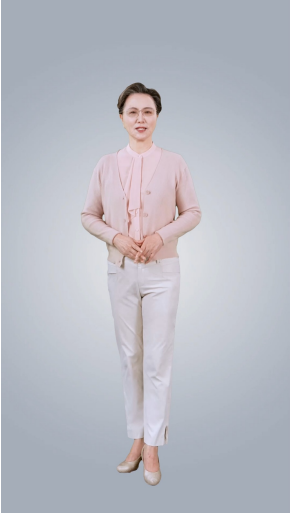
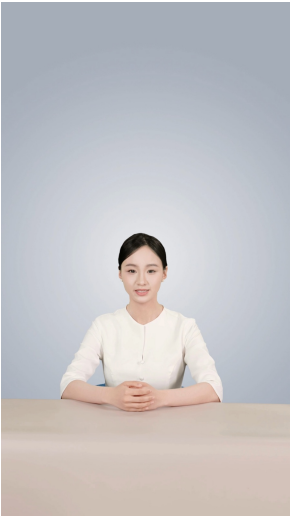
序号	名称	figureid (选择人像)	姿势	背景	性别	人像图	demo视频
1	芝哈	209337	坐姿	实景-横屏	女		芝哈
2	芝哈	209403	坐姿	实景-竖屏	女		芝哈

3	芝哈	211868	站姿	透明	女	非凡人像 	芝哈
4	海霖	209342	坐姿	实景-横屏	男	非凡人像 	海霖
5	海霖	209404	坐姿	实景-竖屏	男	非凡人像 	海霖
6	海霖	211855	站姿	透明	男	非凡人像 	海霖
7	芝怡	209966	坐姿	实景-横屏	女	非凡人像 	芝怡

							
8	芝怡	209967	坐姿	实景-竖屏	女	非凡人像 	芝怡
9	芝怡	211854	站姿	透明	女	非凡人像 	芝怡
10	海昱	210698	坐姿	实景-横屏	男	非凡人像 	海昱
11	海昱	209893	坐姿	实景-竖屏	男	非凡人像 	海昱
						非凡人像 	

12	海昱	211852	站姿	透明	男		海昱
13	芝妍	210697	坐姿	实景-横屏	女		芝妍
14	芝妍	209895	坐姿	实景-竖屏	女		芝妍
15	芝妍	211862	站姿	透明	女		芝妍
16	芝彤	211701	坐姿	实景-横屏	女		芝彤
							

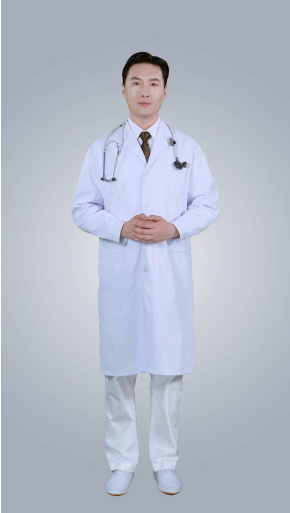
17	芝彤	211700	坐姿	实景-竖屏	女		芝彤
18	芝彤	212414	站姿	透明	女		芝彤
19	慕晴	A2A_V2-muqing	站姿	透明	女		慕晴
20	静怡	A2A_V2-jingyi	站姿	透明	女		静怡
							

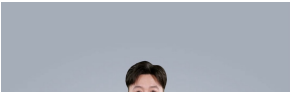
21	璟瑶	A2A_V2-jingyao	站姿	透明	女		璟瑶
22	晚柠	234403	站姿	透明	女		晚柠
23	芊羽	237372	坐姿	透明	女		芊羽
24	芊棠	238260	站姿	透明	女		芊棠
							

25	芊宁	236705	站姿	透明	女		芊宁
26	Amara	238263	站姿	透明	女		Amara
27	初禾	238014	坐姿	透明	女		初禾
28	翊川	238323	站姿	透明	男		翊川
							

29	翊朗	238402	站姿	透明	男		翊朗
30	翊辰	238407	站姿	透明	男		翊辰
31	翊泽	238409	坐姿	透明	男		翊泽
32	Leo	233084	站姿	透明	男		Leo
							

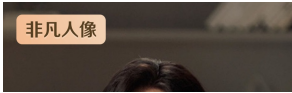
33	Jasper	233056	站姿	透明	男		Jasper
34	景霄	235126	站姿	透明	男		景霄
35	景珩	234731	站姿	透明	男		景珩
36	景舟	234733	站姿	透明	男		景舟
							

37	昭航	227197	站姿	透明	男		昭航
38	昭瑾	234850	站姿	透明	男		昭瑾
39	启铭	238533	坐姿	透明	男		启铭
40	晚意	242844	站姿	透明	女		晚意
							

41	芊梨	238642	站姿	透明	女		芊梨
42	芊舒	238344	站姿	透明	女		芊舒
43	翊轩	238382	坐姿	透明	男		翊轩
44	云澈	238383	站姿	透明	男		云澈
							

45	云琛	238514	站姿	透明	男		云琛
46	云朔	242563	站姿	透明	男		云朔
47	云骁	238641	站姿	透明	男		云骁

🔗 视频场景

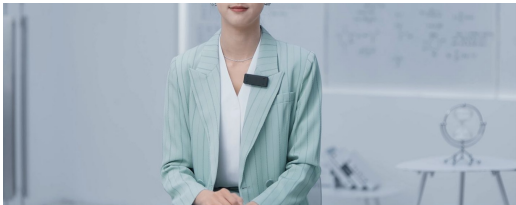
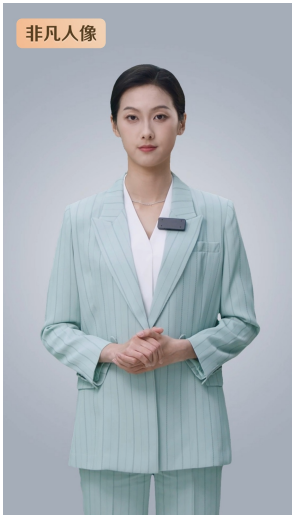
序号	名称	figureid (选择人像)	姿势	背景	性别	人像图	demo视频
1	芝哈	209943	坐姿-正面	实景-横屏	女		芝哈
							

2	芝晗	209965	坐姿	实景-竖屏	女		芝晗
3	芝晗	209979	坐姿	实景-横屏	女		芝晗
4	芝晗	211627	坐姿	实景-竖屏	女		芝晗
5	芝晗	211808	站姿	透明	女		芝晗
6	海霖	209898	坐姿	实景-横屏	男		海霖

7	海霖	209899	坐姿	实景-竖屏	男		海霖
8	海霖	209747	坐姿	实景-横屏	男		海霖
9	海霖	211581	坐姿	实景-竖屏	男		海霖
10	海霖	211809	站姿	透明	男		海霖
11	芝怡	209816	坐姿	实景-横屏	女		芝怡

							
12	芝怡	209963	坐姿	实景-竖屏	女		芝怡
13	芝怡	211562	坐姿-侧面	实景-横屏	女		芝怡
14	芝怡	211624	坐姿-侧面	实景-竖屏	女		芝怡
15	芝怡	211807	站姿	透明	女		芝怡
16	海昱	209886	坐姿-正面	实景-横屏	男		海昱

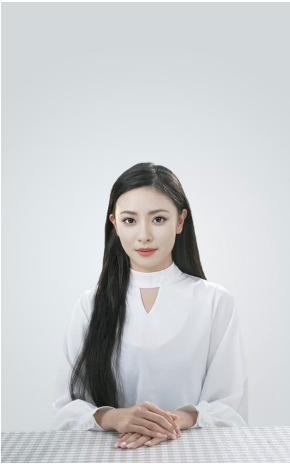
			出	屏			
17	海昱	209897	坐姿-正面	实景-竖屏	男	非凡人像 	海昱
18	海昱	209977	坐姿-侧面	实景-横屏	男	非凡人像 	海昱
19	海昱	211582	坐姿-侧面	实景-竖屏	男	非凡人像 	海昱
20	海昱	211801	站姿	透明	男	非凡人像 	海昱
						非凡人像 	

21	芝妍	209888	坐姿-正面	实景-横屏	女		芝妍
22	芝妍	209896	坐姿-正面	实景-竖屏	女		芝妍
23	芝妍	209978	坐姿-侧面	实景-横屏	女		芝妍
24	芝妍	211621	坐姿-侧面	实景-竖屏	女		芝妍
24	芝妍	211853	站姿	透明	女		芝妍

25	芝彤	209889	坐姿-正面	实景-横屏	女	非凡人像 	芝彤
26	芝彤	209962	坐姿-正面	实景-竖屏	女	非凡人像 	芝彤
27	芝彤	209783	坐姿-侧面	实景-横屏	女	非凡人像 	芝彤
28	芝彤	211623	坐姿-侧面	实景-竖屏	女	非凡人像 	芝彤
28	芝彤	212727	站姿	透明	女	非凡人像 	芝彤

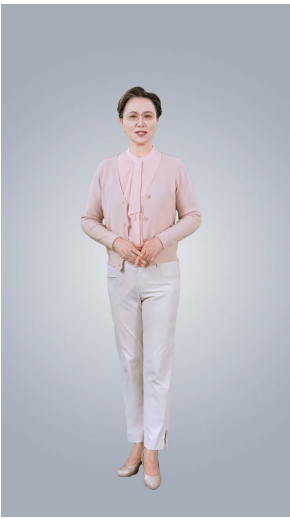
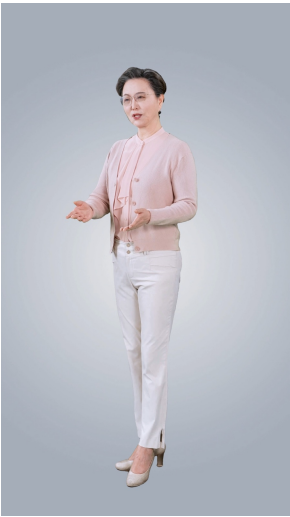
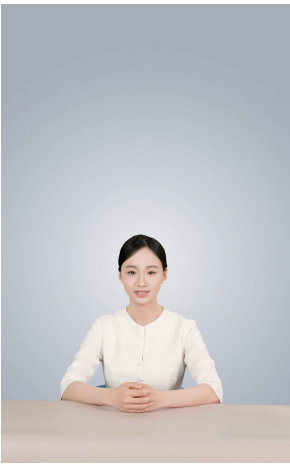
									
29	清馨	1081	站姿	透明	女				清馨
30	清缘	1112	站姿	透明	女				清缘
31	源丞	1090	站姿	透明	男				源丞
32	秋霁	1092	站姿	透明	女				秋霁

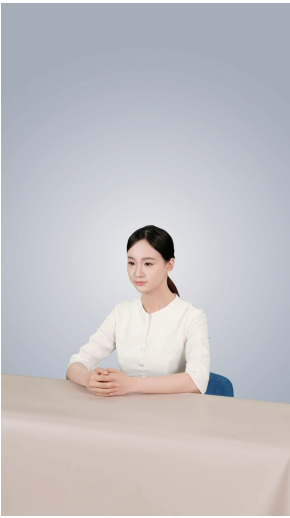
									
33	简兮	1085	站姿	透明	女				简兮
34	沁淑	1087	站姿	透明	女				沁淑
35	沁湘	1088	站姿	透明	女				沁湘
36	步月	1113	站姿	透明	女				步月

									
37	秋霁	1105	坐姿	透明	女				秋霁
38	步月	1107	坐姿	透明	女				步月
39	简兮	1126	坐姿	透明	女				简兮
40	疏影	1109	坐姿	透明	女				疏影

							
41	月言	1192	站姿	透明	女		月言
42	心兰	1207	站姿	透明	女		心兰
43	溪雪	1206	站姿	透明	女		溪雪
44	曼妮	1209	站姿	透明	女		曼妮

									
45	曼琳	1224	坐姿	透明	女				曼琳
46	清妍	1231	站姿	透明	女				清妍
47	彤云	1227	站姿	透明	女				彤云
48	翔宇	1235	站姿	透明	男				翔宇

									
49	飞燕	1244	站姿	透明	女				飞燕
50	晚柠	233129	站姿-正面	透明	女				晚柠
51	晚柠	238017	站姿-侧面	透明	女				晚柠
52	芊羽	238632	坐姿-正面	透明	女				芊羽

									
53	芊羽	238633	坐姿-侧面	透明	女				芊羽
54	芊棠	238635	站姿-正面	透明	女				芊棠
55	芊棠	238636	站姿-侧面	透明	女				芊棠
56	芊宁	238404	站姿-正面	透明	女				芊宁

									
57	芊宁	238406	站姿-侧面	透明	女				芊宁
58	芊梨	238398	站姿-正面	透明	女				芊梨
59	芊梨	238400	站姿-侧面	透明	女				芊梨
60	翊辰	238420	站姿-正面	透明	男				翊辰

61	翊辰	238397	站姿-侧面	透明	男				翊辰
62	Leo	236252	站姿-正面	透明	男				Leo
63	Leo	236611	站姿-侧面	透明	男				Leo
64	Jasper	236311	站姿-正面	透明	男				Jasper

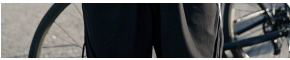
									
65	Jasper	236678	站姿-侧面	透明	男				Jasper
66	景霄	233057	站姿-正面	透明	男				景霄
67	景霄	238097	站姿-侧面	透明	男				景霄
68	景珩	236519	站姿-正面	透明	男				景珩

									
69	景珩	238120	站姿-侧面	透明	男				景珩
70	景舟	238060	站姿-正面	透明	男				景舟
71	景舟	238129	站姿-侧面	透明	男				景舟
72	启铭	238525	坐姿-正面	透明	男				启铭

									
73	启铭	238526	坐姿-侧面	透明	男				启铭
74	承屿	234270	站姿	实景	男				承屿
75	砚博	234331	站姿	实景	男				砚博
76	屹阳	234332	站姿	实景	男				屹阳

								
77	景钧	234441	站姿	实景	男	生成式人像 	景钧	
78	茉斐	234335	站姿	实景	女	生成式人像 	茉斐	
79	砚秋	234336	站姿	实景	女	生成式人像 	砚秋	
80	砚朗	234337	站姿	实景	男	生成式人像 	砚朗	

								
81	若衡	234338	站姿	实景	女	生成式人像 	若衡	
82	昭泽	234340	站姿	实景	男	生成式人像 	昭泽	
83	南音	234393	站姿	实景	女	生成式人像 	南音	
84	知越	235028	站姿	实景	男	生成式人像 	知越	

								
85	夏柠	235075	站姿	实景	女	生成式人像 	夏柠	
86	沐云	234452	站姿	实景	男	生成式人像 	沐云	
87	沐勋	234729	站姿	实景	男	生成式人像 	沐勋	
88	沐音	234793	站姿	实景	男	生成式人像 	沐音	

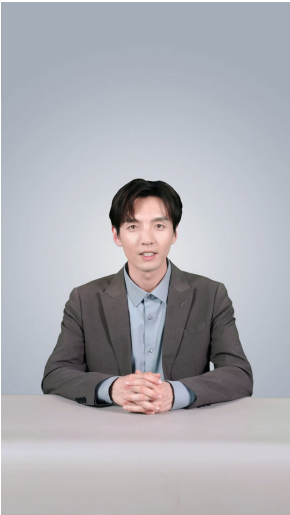
							
89	沐白	234724	站姿	实景	男	生成式人像 	沐白
90	沐逸	234716	站姿	实景	男	生成式人像 	沐逸
91	初瑶	234278	站姿	实景	女	生成式人像 	初瑶
92	屹尘	234932	站姿	实景	男	生成式人像 	屹尘

							
93	Rock	234965	坐姿	实景	男	生成式人像 	Rock
94	昭航	238009	站姿-正面	透明	男		昭航
95	昭航	238305	站姿-侧面	透明	男		昭航
96	昭瑾	236251	站姿-正面	透明	男		昭瑾

									
97	昭瑾	238646	站姿-侧面	透明	男				昭瑾
98	晚意	244790	站姿-正面	透明	女				晚意
99	晚意	236735	站姿-侧面	透明	女				晚意
100	芊舒	243226	站姿-正面	透明	女				芊舒

									
101	芊舒	243228	站姿-侧面	透明	女				芊舒
102	Amara	238334	站姿-正面	透明	女				Amara
103	Amara	244519	站姿-侧面	透明	女				Amara
104	初禾	238362	坐姿-正面	透明	女				初禾

									
105	翊川	238794	站姿-正面	透明	男				翊川
106	翊川	238375	站姿-侧面	透明	男				翊川
107	翊朗	238731	站姿-正面	透明	男				翊朗
108	翊朗	238405	站姿-侧面	透明	男				翊朗

									
109	翊泽	238802	坐姿-正面	透明	男				翊泽
110	翊泽	238698	坐姿-侧面	透明	男				翊泽
111	翊轩	238684	坐姿-正面	透明	男				翊轩
112	翊轩	238685	坐姿-侧面	透明	男				翊轩

									
113	云澈	238686	站姿-正面	透明	男				云澈
114	云澈	238687	站姿-侧面	透明	男				云澈
115	云琛	238688	站姿-正面	透明	男				云琛
116	云琛	238689	站姿-侧面	透明	男				云琛

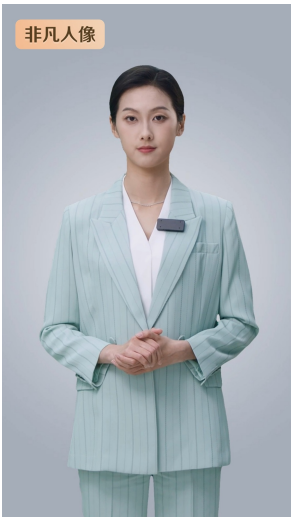
									
117	云朔	243222	站姿-正面	透明	男				云朔
118	云朔	243223	站姿-侧面	透明	男				云朔
119	云骁	238432	站姿-正面	透明	男				云骁
120	云骁	238433	站姿-侧面	透明	男				云骁

--	--	--	--	--	--	--	--

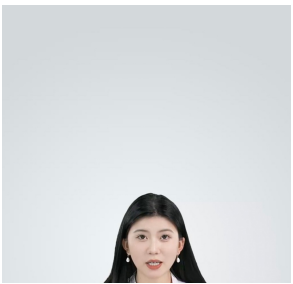
直播场景

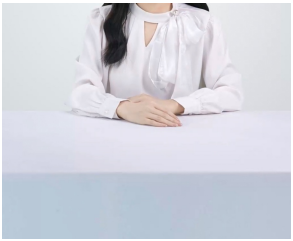
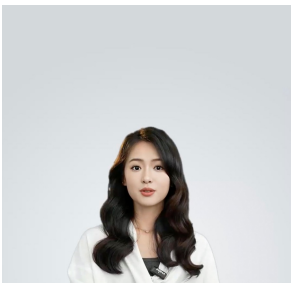
序号	名称	figureid（选择人像）	姿势	背景	性别	人像图	demo视频
1	芝哈	209403	坐姿	实景-竖屏	女	非凡人像 	芝哈
2	芝哈	211868	站姿	透明	女	非凡人像 	芝哈
3	海霖	209404	坐姿	实景-竖屏	男	非凡人像 	海霖
4	海霖	211855	站姿	透明	男	非凡人像 	海霖

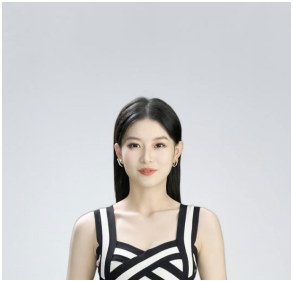
							
5	芝怡	209967	坐姿	实景-竖屏	女	非凡人像 	芝怡
6	芝怡	211854	站姿	透明	女	非凡人像 	芝怡
7	海昱	209893	坐姿	实景-竖屏	男	非凡人像 	海昱
8	海昱	211852	站姿	透明	男	非凡人像 	海昱

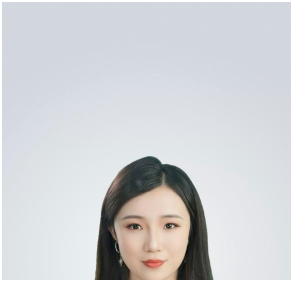
							
9	芝妍	209895	坐姿	实景-竖屏	女	非凡人像 	芝妍
38	芝妍	211853	站姿	透明	女	非凡人像 	芝妍
10	芝彤	211700	坐姿	实景-竖屏	女	非凡人像 	芝彤
10	芝彤	212727	站姿	透明	女	非凡人像 	芝彤

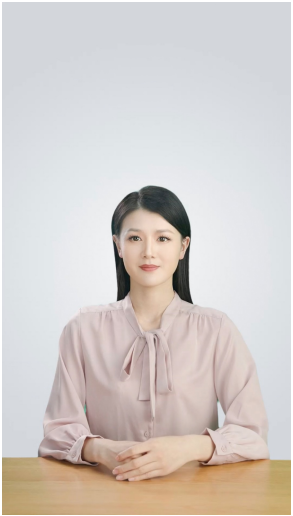
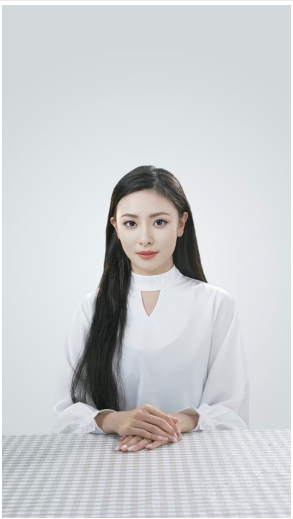
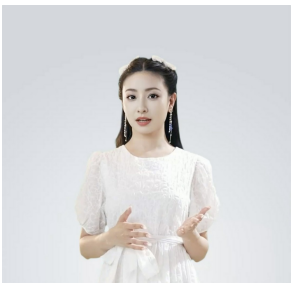
							
11	璟瑶	1104	站姿	透明	女		璟瑶站姿
12	璟瑶	1005	坐姿	透明	女		璟瑶坐姿
13	语欣	1121	坐姿	透明	女		语欣坐姿
14	佳钰	1008	站姿	透明	女		佳钰站姿

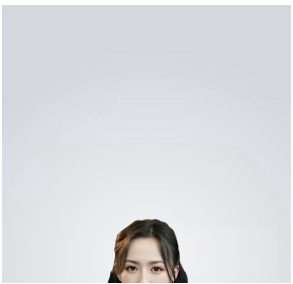
							
15	月婵	1119	站姿	透明	女		月婵站姿
16	月婵	1124	坐姿	透明	女		月婵坐姿
17	语桐	1130	站姿	透明	女		语桐站姿
18	慕晴	1138	坐姿	透明	女		慕晴坐姿

							
19	龚星	1093	坐姿	透明	女		龚星
20	博菊	1098	坐姿	透明	女		博菊
21	刘畅	1094	坐姿	透明	女		刘畅
22	清馨	1081	站姿	透明	女		清馨

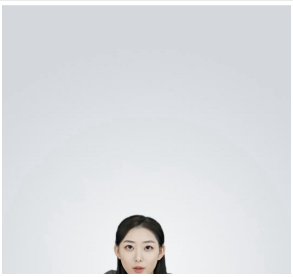
							
23	清缘	1112	站姿	透明	女		清缘
24	源丞	1090	站姿	透明	男		源丞
25	秋霁	1092	站姿	透明	女		秋霁
26	简兮	1085	站姿	透明	女		简兮

							
27	沁淑	1087	站姿	透明	女		沁淑
28	沁湘	1088	站姿	透明	女		沁湘
29	步月	1113	站姿	透明	女		步月
30	秋霁	1105	坐姿	透明	女		秋霁

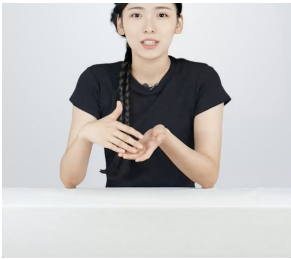
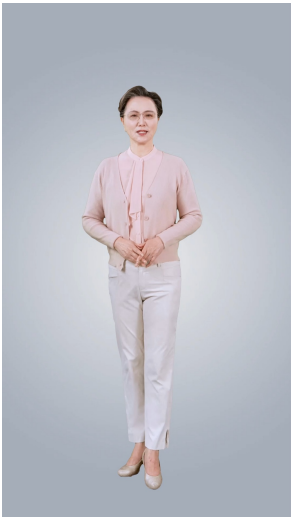
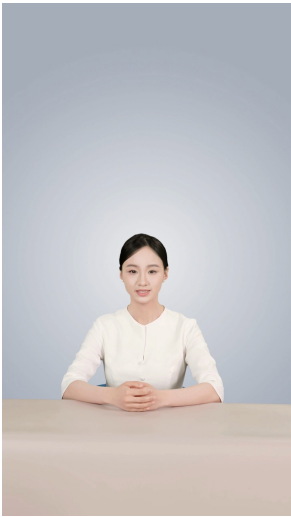
							
31	步月	1107	坐姿	透明	女		步月
32	简兮	1126	坐姿	透明	女		简兮
33	疏影	1109	坐姿	透明	女		疏影
34	月言	1192	站姿	透明	女		月言

							
35	心兰	1207	站姿	透明	女		心兰
36	溪雪	1206	站姿	透明	女		溪雪
37	曼妮	1209	站姿	透明	女		曼妮
38	曼琳	1224	坐姿	透明	女		曼琳

							
39	清妍	1231	站姿	透明	女		清妍
40	彤云	1227	站姿	透明	女		彤云
41	翔宇	1235	站姿	透明	男		翔宇
42	飞燕	1244	站姿	透明	女		飞燕

							
43	静怡	1181	站姿	透明	女		静怡
44	璟妍	1183	站姿	透明	女		璟妍
45	董林	1213	坐姿	透明	男		董林
46	莉白	1214	坐姿	透明	女		莉白

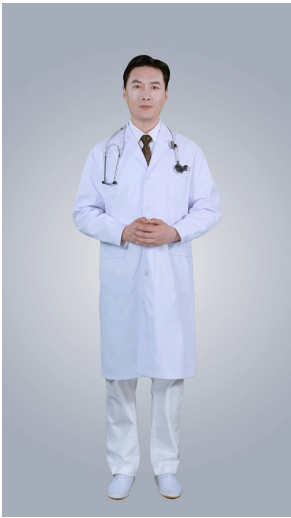
							
47	赵岐	1216	坐姿	透明	女		赵岐
48	代乐	1221	坐姿	透明	女		代乐
49	孙昕	1222	坐姿	透明	女		孙昕
50	洪娜	1223	坐姿	透明	女		洪娜

							
51	晚柠	234403	站姿	透明	女		晚柠
52	芊羽	237372	坐姿	透明	女		芊羽
53	芊棠	238260	站姿	透明	女		芊棠
54	芷宁	236705	站姿	透明	女		芷宁

	丁	丁	238100	站姿	透明	女		丁
55	Amara		238263	站姿	透明	女		Amara
56	初禾		238014	坐姿	透明	女		初禾
57	翊川		238323	站姿	透明	男		翊川
58	翊阳		238402	站姿	透明	男		翊阳

	翊辰	238402	站姿	透明	男		翊辰
59	翊辰	238407	站姿	透明	男		翊辰
60	翊泽	238409	坐姿	透明	男		翊泽
61	Leo	233084	站姿	透明	男		Leo
62	lasner	233056	站姿	透明	男		lasner

62	景霄	235126	站姿	透明	男		景霄
63	景霄	235126	站姿	透明	男		景霄
64	景珩	234731	站姿	透明	男		景珩
65	景舟	234733	站姿	透明	男		景舟
66	昭皓	227197	站姿	透明	男		昭皓

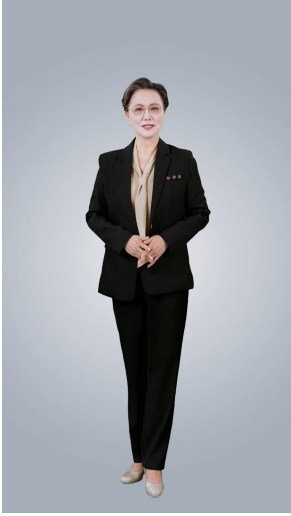
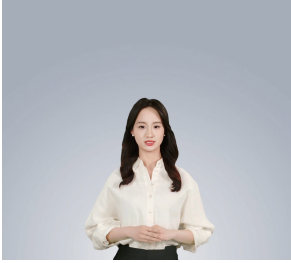
	昭瑾	2341201	站姿	透明	男		昭瑾
67	昭瑾	234850	站姿	透明	男		昭瑾
68	启铭	238533	坐姿	透明	男		启铭
69	承屿	234270	站姿	实景	男		承屿
70	碑博	234331	站姿	实景	男		碑博

	屹阳	234331	站姿	实景	男		屹阳
71	屹阳	234332	站姿	实景	男	生成式人像 	屹阳
72	景钧	234441	站姿	实景	男	生成式人像 	景钧
73	茉斐	234335	站姿	实景	女	生成式人像 	茉斐
74	砚秋	234336	站姿	实景	女	生成式人像 	砚秋

	75	砚朗	234337	站姿	实景	男		砚朗
	76	若衡	234338	站姿	实景	女	生成式人像 	砚朗
	77	昭泽	234340	站姿	实景	男	生成式人像 	若衡
	78	南音	234393	站姿	实景	女	生成式人像 	昭泽
							生成式人像 	南音

	知越	2345000	站姿	实景	女		知越
79	知越	235028	站姿	实景	男	生成式人像 	知越
80	夏柠	235075	站姿	实景	女	生成式人像 	夏柠
81	沐云	234452	站姿	实景	男	生成式人像 	沐云
82	沐勐	234729	站姿	实景	男	生成式人像 	沐勐

							
83	沐音	234793	站姿	实景	男	生成式人像 	沐音
84	沐白	234724	站姿	实景	男	生成式人像 	沐白
85	沐逸	234716	站姿	实景	男	生成式人像 	沐逸
86	初瑶	234278	站姿	实景	女	生成式人像 	初瑶

							
87	屹尘	234932	站姿	实景	男	生成式人像 	屹尘
88	Rock	234965	坐姿	实景	男	生成式人像 	Rock
89	晚意	242844	站姿	透明	女		晚意
90	芊梨	238642	站姿	透明	女		芊梨

							↑
91	芊舒	238344	站姿	透明	女		芊舒
92	翊轩	238382	坐姿	透明	男		翊轩
93	云澈	238383	站姿	透明	男		云澈
94	云琛	238514	站姿	透明	男		云琛

							
95	云朔	242563	站姿	透明	男		云朔
96	云骁	238641	站姿	透明	男		云骁

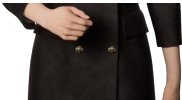
3D数字人

2025年6月起新增人像将仅在配置台更新，数字人figureid请从 <https://xiling.cloud.baidu.com/open/config> 获取。

形象 - figure id

形象适用范围：云渲染交互组件、视频组件

序号	名称	figureid (人像 id)	人像图
1	昭宁 适用交互组件	A2A_V2 -fig- reyuyti gtk7igh h	

				
2	昭宁 适用视频组件 (2k)	A2A_V2 -fig- rexqufn 4wxwgr 6ne		
3	星瑶 适用交互组件	A2A_V2 -fig- reyu7d 545k1d kdcn		
4	星瑶 适用视频组件 (2k)	A2A_V2 -fig- rexqu1 94cr14i 79b		
5	云晞 适用交互组件	A2A_V2 -fig- reyur71 quhggp wv5		
				

6	云晞 适用视 频组件 (2k)	A2A_V2 -fig- rexu9s uf82wb bzsbs	
7	昭宁 适用交 互组件	A2A_V2 -fig- rd4mkz 79idnh ydu0	
8	昭宁 适用视 频组件	A2A_V2 -fig- regmjs wrdw0i 8kqt	

9	语棠 适用交互组件	A2A_V2 -fig- rd4sc8 56zcwi d5ak	
10	语棠 适用视频组件	A2A_V2 -fig- remtdp 69y7x1 s9ez	
11	乐然 适用交互组件	A2A_V2 -fig- rd4st6s sta2gvv	

		8y	
12	乐然 适用视频组件	A2A_V2 -fig- remqxgt vy7jw67 pg	
13	星瑶 适用交互组件	A2A_V2 -fig- rd4sjmj mqy42v ze7	

14	星瑶 适用视频组件	A2A_V2 -fig-remtjxqvi3db7vci	
15	婉宁 适用交互组件	A2A_V2 -fig-rd4vgh6qtm6pe1iq	
16	婉宁 适用视频组件	A2A_V2 -fig-remsrdxztkn2c71~	

		lp	
17	舒颜 适用交互组件	A2A_V2 -fig- rd4spx p648p3 965h	
18	舒颜 适用视频组件	A2A_V2 -fig- remvf7c fum9tm b7f	
			

19	甜夏 适用交互组件	A2A_V2 -fig- rd5rz8f 1368gf maw	
20	甜夏 适用视频组件	A2A_V2 -fig- reft6ss bx6n48 srq	
21	晓棠 适用交互组件	A2A_V2 -fig- rd4t9kd ktapa9r r2	

				
22	晓棠 适用视 频组件	A2A_V2 -fig- remtiuk 55uzrw 48j		
23	云晞 适用交 互组件	A2A_V2 -fig- rd4tjyy4 vczedfh a		
				

24	云晞 适用视频组件	A2A_V2 -fig-remsr3ab3ikrm114	
25	清妍 适用交互组件	A2A_V2 -fig-rd4sgj8ryptwq06e	
26	清妍 适用视频组件	A2A_V2 -fig-remr35w1edtf7djr	

				
27	依依-新春 适用交互组件	A2A_V2 -fig- rarusez s1p0y5 x6s		
28	依依-新春 适用视频组件	A2A_V2 -fig- rasuhd u9x4an 4fxe		
				

29	苏瑾-新春 适用交互组件	A2A_V2 -fig- raqqhrb 2y7vi19 z1		
30	苏瑾-新春 适用视频组件	A2A_V2 -fig- rarvdyy e4q1ek uum		
31	思远-新春 适用交互组件	A2A_V2 -fig- rart5j3p 2i1bkzx n		

				
32	美涵·新春 适用交互组件	A2A_V2 -fig- rasr44u r7bt7is e1		
33	美涵	A2A_V2 - 3to2_m eihan		
				

34	依依	A2A_V2 - 3to2_yi yi	
35	可欣	A2A_V2 - 3to2_k exin	
36	乐晴	A2A_V2 - 3to2_le qing	

				
37	苏瑾	A2A_V2 - 3to2_s ujin		
38	馨怡	A2A_V2 - 3to2_xi nyi		
				

39	子悦	A2A_V2 - 3to2_zi yue	
40	鑫睿	A2A_V2 - 3to2_xi nrui	
41	思瑾	A2A_V2 - 3to2_sij in	

42	思远	A2A_V2 - 3to2_si yuan	
43	慧雅	A2A_V2 - 3to2_h uiya	

2D精品数字人

2025年6月起新增人像将仅在配置台更新，数字人figureid请从 <https://xiling.cloud.baidu.com/open/config> 获取。

1 机位 - camera id

16:9 横屏比例

- 半身：0
- 全身：1

9:16 竖屏比例

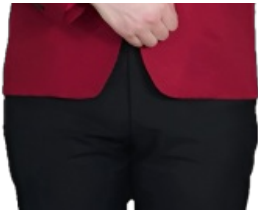
- 半身：2
- 全身：3

2 形象 - figure id

形象适用范围：云渲染交互组件、视频组件

序号	名称	figureid（选择人像）	人像图
1	梓欣	A2A_V2-xinxin	
2	筱萱	A2A_V2-xixi	
			

3	乔雅	A2A_V2-xiaomeng2	
4	嘉睿	A2A_V2-aning	
5	嘉霖	A2A_V2-aning_red	

			
6	纪楚	A2A_V2-gaoming	

端渲染数字人

🔗 端渲染2d形象 - 人像包

形象适用范围：端染交互组件-2d数字人

序号	名称	形象包	预览图
1	璟馨-站姿	https://meta-human-editor-prd.bj.bcebos.com/open-api/jingxin-stand0328.tar.gz	
2	慕晴-站姿	https://meta-human-editor-prd.bj.bcebos.com/open-api/muqing-stand0328.tar.gz	
3	晓峰-站姿	https://meta-human-editor-prd.bj.bcebos.com/open-api/xiaofeng-stand0328.tar.gz	

端渲染3d形象 - 人像包

形象适用范围：端染交互组件-3d数字人

序号	名称	类型	h5资产包	双端资产包（Android、iOS）	预览图（示意图为固定搭配，不支持混搭）

1	yuki	卡通	https://meta-human-editor-prd.bj.bcebos.com/open-api/yuki_h5_250206.zip	https://meta-human-editor-prd.bj.bcebos.com/open-api/yuki_client_250117.zip	
2	judy	写实	https://meta-human-editor-prd.cdn.bcebos.com/open-api/Judy_h5_250212.zip	-	
3	coco	写实	https://meta-human-editor-prd.cdn.bcebos.com/open-api/coco_h5_250305.zip	https://meta-human-editor-prd.bj.bcebos.com/open-api/case-coco-20250401.zip	

					
4	charis	写实	https://digital-human-js-cdn.cdn.bcebos.com/paas-demo/charis_h5-20250206.zip	-	 

公共音色库

🔗 多语言音色

* 支持方言及语种共计23种:Chinese（中文）；Chinese,Yue（粤语）；English（英语）；Russian（俄语）；Spanish（西班牙语）；French（法语）；Portuguese（葡萄牙语）；German（德语）；Turkish（土耳其语）；Dutch（荷兰语）；Ukrainian（乌克兰语）；Vietnamese（越南语）；Indonesian（印尼语）；Japanese（日语）；Italian（意大利语）；Korean（韩语）；Thai（泰语）；Polish（波兰语）；Romanian（罗马尼亚语）；Greek（希腊语）；Czech（捷克语）；Finnish（芬兰语）；Hindi（印地语）

下表中每个音色均支持多语言，音色效果可通过 [试听文件](#) 体验。

序号	名称	音色ID	性别	年龄	风格	特征
1	内敛闷声叙述者	LITE_Arnold	男	25-35岁	亲和陪伴	低沉磁性
2	沉静知性叙述者	LITE_Attractive_Girl	女	25-35岁	亲和陪伴	低沉磁性
3	端庄专业女主播	LITE_audiobook_female_1	女	25-35岁	权威专业	温润柔和
4	成熟可靠女主播	LITE_audiobook_female_2	女	25-35岁	权威专业	温润柔和
5	清朗专业男主播	LITE_audiobook_male_1	男	18-24岁	权威专业	温润柔和
6	抑扬顿挫说书人	LITE_audiobook_male_2	男	36-50岁	权威专业	温润柔和
7	声情并茂解说员	LITE_badao_shaoye	男	25-35岁	权威专业	温润柔和
8	柔情细语小哥哥	LITE_bingjiao_didi	男	18-24岁	亲和陪伴	温润柔和
9	轻盈童声播报员	LITE_cartoon_pig	女	18岁以下	权威专业	清透明亮
10	轻声低哑小姐姐	LITE_Charming_Lady	女	18-24岁	亲和陪伴	低沉磁性
11	悬疑神秘叙述者	LITE_Charming_Santa	男	25-35岁	权威专业	低沉磁性
12	温和理性小哥哥	LITE_chunzhen_xuedi	男	18-24岁	权威专业	温润柔和
13	乖巧聪明小男孩	LITE_clever_boy	男	18岁以下	亲和陪伴	温润柔和
14	可爱q弹小男孩	LITE_cute_boy	男	18岁以下	亲和陪伴	温润柔和
15	尖锐活泼卡通音	LITE_Cute_Elf	女	18岁以下	活力激情	清透明亮
16	干练可靠研究员	LITE_danya_xuejie	女	25-35岁	权威专业	低沉磁性
17	欢快洪亮卡通音	LITE_diadia_xuemei	女	18岁以下	活力激情	清透明亮
18	成熟稳重女主播	LITE_female-chengshu	女	25-35岁	权威专业	低沉磁性
19	专业干练女主播	LITE_female-chengshu-jingpin	女	25-35岁	权威专业	低沉磁性
20	温柔细腻小姐姐	LITE_female-shaonv	女	18-24岁	亲和陪伴	温润柔和
21	亲切温和小姐姐	LITE_female-shaonv-jingpin	女	18-24岁	亲和陪伴	温润柔和
22	轻声细语小姐姐	LITE_female-tianmei	女	18-24岁	亲和陪伴	温润柔和
23	文静内敛小姐姐	LITE_female-tianmei-jingpin	女	18-24岁	亲和陪伴	温润柔和
24	慵懒松弛大姐姐	LITE_female-yujie	女	25-35岁	亲和陪伴	温润柔和
25	温柔慵懒大姐姐	LITE_female-yujie-jingpin	女	25-35岁	亲和陪伴	温润柔和
26	热情爽朗讲解员	LITE_Grinch	男	36-50岁	活力激情	清透明亮
27	活力热情小哥哥	LITE_junlang_nanyou	男	18-24岁	活力激情	清透明亮
28	成熟稳重小哥哥	LITE_lengdan_xiongzhang	男	18-24岁	权威专业	低沉磁性
29	稚嫩纯真小女孩	LITE_lovely_girl	女	18岁以下	亲和陪伴	清透明亮
30	亲切质朴讲解员	LITE_male-qn-badao	男	18-24岁	权威专业	温润柔和
31	冷静理智小哥哥	LITE_male-qn-badao-jingpin	男	18-24岁	权威专业	温润柔和
32	纯真温和小哥哥	LITE_male-qn-daxuesheng	男	18-24岁	亲和陪伴	温润柔和
33	文静理智小哥哥	LITE_male-qn-daxuesheng-jingpin	男	18-24岁	亲和陪伴	温润柔和
34	干练可靠讲解员	LITE_male-qn-jingying	男	25-35岁	活力激情	低沉磁性
35	稳重可靠讲解员	LITE_male-qn-jingying-jingpin	男	25-35岁	权威专业	低沉磁性
36	阳光温和小哥哥	LITE_male-qn-qingse	男	18-24岁	亲和陪伴	温润柔和
37	青涩随和小哥哥	LITE_male-qn-qingse-jingpin	男	18-24岁	亲和陪伴	温润柔和
38	成熟干练女主播	LITE_presenter_female	女	25-35岁	权威专业	温润柔和

39	成熟干练男主播	LITE_presenter_male	男	25-35岁	权威专业	温润柔和
40	可爱稚气小主播	LITE_qiaopi_mengmei	女	18岁以下	权威专业	清透明亮
41	温厚沉稳解说者	LITE_Rudolph	女	25-35岁	权威专业	温润柔和
42	声情并茂男主播	LITE_Santa_Claus	男	25-35岁	活力激情	温润柔和
43	温柔甜美讲解员	LITE_Serene_Woman	女	18-24岁	权威专业	清透明亮
44	慵懒温柔小姐姐	LITE_Sweet_Girl	女	18-24岁	亲和陪伴	温润柔和
45	甜美亲切小姐姐	LITE_tianxin_xiaoling	女	18-24岁	亲和陪伴	温润柔和
46	亲和温柔解说员	LITE_wumei_yujie	女	18-24岁	亲和陪伴	温润柔和

🔊 女声音色

序号	音色名称	音色ID	性别	风格	适用场景	试听音频文件
1	度禧禧	CAP_4146	女	温柔甜美	客服助手	度禧禧
2	度小柔	CAP_6567	女	知性大方	客服助手	度小柔
3	度涵竹	CAP_4189	女	自然生动	广告营销	度涵竹
4	度嫣然	CAP_4194	女	温柔可爱	客服助手	度嫣然
5	度清影	CAP_4196	女	甜美可爱	广告营销	度清影
6	度沁遥	CAP_4197	女	温柔知性	客服助手	度沁遥
7	度小夏	5132	女	知性大方	客服助理/新闻主持/广告营销/教师讲座	度小夏
8	度小雯	4100	女	元气活力	客服助理/广告营销	度小雯
9	度小希	5116	女	元气活力	广告营销	度小希
10	度常盈	5147	女	亲和力强	客服助理/广告营销	度常盈
11	度灵儿	4105	女	权威靠谱	广告营销	度灵儿
12	度小乔	4117	女	知性大方	广告营销	度小乔
13	度小鹿	4118	女	元气活力	客服助理/广告营销	度小鹿
14	度小语	4000	女	权威靠谱	广告营销	度小语
15	度常悦	5153	女	知性大方	客服助理/广告营销	度常悦
16	度仔仔	5137	女	元气活力	广告营销	度仔仔
17	西贝	4125	女	激情饱满	广告营销	西贝
18	专业靠谱爽朗女	7011_moxingxiaoxiao_16k	女	专业娴熟/沉稳冷静/激情饱满	广告营销/电商直播	专业靠谱爽朗女
19	热情悦耳女主播	7011_moxingkangxi_16k	女	元气活力/权威靠谱/激情饱满	广告营销/电商直播	热情悦耳女主播
20	自信活泼小姐姐	7011_moxinghuanhuan_16k	女	元气活力/权威靠谱/沉稳冷静	客服助理/广告营销/电商直播	自信活泼小姐姐
21	自信活泼大姐姐	7011_vc0057_16k	女	元气活力/权威靠谱/沉稳冷静	广告营销/电商直播	自信活泼大姐姐
22	自然朴实小妹妹	7011_vc0020_16k	女	专业娴熟/亲和力强/权威靠谱	客服助理/广告营销/电商直播	自然朴实小妹妹
	专注真诚大			专业娴熟/亲和力强/		专注真诚大

23	姐姐	7011_vc0053_16k	女	权威靠谱	广告营销/电商直播	姐姐
24	职业霸气御姐	7011_vc0033_16k	女	专业娴熟/权威靠谱/沉稳冷静	客服助理/新闻主持/广告营销/教师讲座/电商直播	职业霸气御姐
25	知性优雅叙事女声	7011_vc0019_16k	女	专业娴熟/亲和力强/沉稳冷静	客服助理/新闻主持/广告营销/教师讲座/电商直播	知性优雅叙事女声
26	幽默东北大妹子	7011_vc0048_16k	女	亲和力强/权威靠谱/激情饱满	广告营销/电商直播	幽默东北大妹子
27	优雅沉稳女主播	7011_vc0022_16k	女	亲和力强/元气活力/沉稳冷静	新闻主持/广告营销/教师讲座/电商直播	优雅沉稳女主播
28	音色纯净悠扬声	7011_vc0140_16k	女	专业娴熟/亲和力强/权威靠谱	广告营销/电商直播	音色纯净悠扬声
29	严谨细致推销员	7011_vc0002_16k	女	专业娴熟/权威靠谱/沉稳冷静	新闻主持/广告营销/教师讲座/电商直播	严谨细致推销员
30	学识渊博知性女	7011_vc0127_16k	女	亲和力强/权威靠谱/沉稳冷静	客服助理/新闻主持/广告营销/教师讲座/电商直播	学识渊博知性女
31	性情温和女主播	7011_vc0080_16k	女	亲和力强/元气活力/权威靠谱	客服助理/新闻主持/广告营销/教师讲座/电商直播	性情温和女主播
32	细心果断女主播	7011_vc0039_16k	女	元气活力/权威靠谱/激情饱满	新闻主持/广告营销/教师讲座/电商直播	细心果断女主播
33	温文尔雅女生	7011_vc0023_16k	女	专业娴熟/权威靠谱/沉稳冷静	新闻主持/广告营销/教师讲座/电商直播	温文尔雅女生
34	温婉秀丽靓女	7011_vc0130_16k	女	亲和力强/幽默有趣/权威靠谱	客服助理/广告营销/电商直播	温婉秀丽靓女
35	温婉细致淑女	7011_vc0117_16k	女	专业娴熟/元气活力/权威靠谱	广告营销/电商直播	温婉细致淑女
36	温婉随和大姐姐	7011_vc0126_16k	女	专业娴熟/权威靠谱/沉稳冷静	广告营销/电商直播	温婉随和大姐姐
37	温柔真诚大姐姐	7011_vc0056_16k	女	亲和力强/权威靠谱/沉稳冷静	广告营销/电商直播	温柔真诚大姐姐
38	温柔体贴靓女	7011_vc0031_16k	女	亲和力强/元气活力/沉稳冷静	新闻主持/广告营销/教师讲座/电商直播	温柔体贴靓女
39	温柔体贴大姐姐	7011_vc0013_16k	女	专业娴熟/亲和力强/激情饱满	新闻主持/广告营销/教师讲座/电商直播	温柔体贴大姐姐
40	温柔亲和女主播	7011_vc0114_16k	女	元气活力/权威靠谱/激情饱满	广告营销/电商直播	温柔亲和女主播
41	温柔磁性女声	7011_vc0070_16k	女	专业娴熟/亲和力强/元气活力	客服助理/广告营销/电商直播	温柔磁性女声
42	恬静文雅气质女	7011_vc0119_16k	女	亲和力强/元气活力/激情饱满	广告营销/电商直播	恬静文雅气质女
43	恬静温柔小姐姐	7011_vc0112_16k	女	元气活力/幽默有趣/激情饱满	客服助理/广告营销/电商直播	恬静温柔小姐姐
44	思维敏锐女主播	7011_vc0032_16k	女	专业娴熟/亲和力强/元气活力	新闻主持/广告营销/教师讲座/电商直播	思维敏锐女主播
45	思维敏捷讲	7011_vc0108_16k	女	专业娴熟/亲和力强/	广告营销/电商直播	思维敏捷讲

45	解员	7011_vc0100_16k	女	元气活力	客服助理/电商直播	解员
46	爽朗外向大姐姐	7011_vc0118_16k	女	幽默有趣/权威靠谱/ 沉稳冷静	客服助理/新闻主持/广告营销/教师 讲座/电商直播	爽朗外向大姐姐
47	爽朗东北靓女	7011_vc0027_16k	女	专业娴熟/元气活力/ 沉稳冷静	新闻主持/广告营销/教师讲座/电商 直播	爽朗东北靓女
48	爽快果断女主播	7011_vc0045_16k	女	亲和力强/元气活力/ 幽默有趣	广告营销/电商直播	爽快果断女主播
49	爽快东北大姐	7011_vc0017_16k	女	元气活力/幽默有趣/ 激情饱满	广告营销/电商直播	爽快东北大姐
50	帅气中性女声	7011_vc0102_16k	女	亲和力强/元气活力/ 激情饱满	新闻主持/广告营销/教师讲座/电商 直播	帅气中性女声
51	善良朴实大姐	7011_vc0115_16k	女	专业娴熟/元气活力/ 幽默有趣	新闻主持/广告营销/教师讲座/电商 直播	善良朴实大姐
52	善解人意女主播	7011_vc0009_16k	女	专业娴熟/亲和力强/ 沉稳冷静	新闻主持/广告营销/教师讲座/电商 直播	善解人意女主播
53	善解人意大姐姐	7011_vc0007_16k	女	专业娴熟/亲和力强/ 沉稳冷静	客服助理/新闻主持/广告营销/教师 讲座/电商直播	善解人意大姐姐
54	热情周到闺蜜	7011_vc0125_16k	女	亲和力强/幽默有趣/ 沉稳冷静	广告营销/电商直播	热情周到闺蜜
55	热情真实女主播	7011_vc0052_16k	女	元气活力/权威靠谱/ 激情饱满	广告营销/电商直播	热情真实女主播
56	热情洋溢大姐大	7011_vc0005_16k	女	亲和力强/元气活力/ 激情饱满	广告营销/电商直播	热情洋溢大姐大
57	热情开朗女孩	7011_vc0004_16k	女	专业娴熟/元气活力/ 激情饱满	广告营销/电商直播	热情开朗女孩
58	清爽洒脱大姐姐	7011_vc0024_16k	女	专业娴熟/亲和力强/ 元气活力	新闻主持/广告营销/教师讲座/电商 直播	清爽洒脱大姐姐
59	清脆婉转动听音	7011_vc0138_16k	女	元气活力/权威靠谱/ 激情饱满	客服助理/广告营销/电商直播	清脆婉转动听音
60	勤勉活力女销冠	7011_vc0120_16k	女	亲和力强/权威靠谱/ 沉稳冷静	客服助理/广告营销/电商直播	勤勉活力女销冠
61	亲切随和女声	7011_vc0090_16k	女	元气活力/权威靠谱/ 激情饱满	广告营销/电商直播	亲切随和女声
62	亲切随和讲解员	7011_vc0110_16k	女	元气活力/幽默有趣/ 激情饱满	广告营销/电商直播	亲切随和讲解员
63	朴实大方邻家小妹	7011_vc0051_16k	女	专业娴熟/亲和力强/ 沉稳冷静	客服助理/广告营销/电商直播	朴实大方邻家小妹
64	男友力爆表酷妹	7011_vc0095_16k	女	专业娴熟/亲和力强/ 沉稳冷静	新闻主持/广告营销/教师讲座/电商 直播	男友力爆表酷妹
65	美妙动听女主播	7011_vc0136_16k	女	亲和力强/元气活力/ 权威靠谱	新闻主持/广告营销/教师讲座/电商 直播	美妙动听女主播
66	利落干脆小姐姐	7011_vc0096_16k	女	元气活力/幽默有趣/ 激情饱满	广告营销/电商直播	利落干脆小姐姐
67	乐观向上小姐姐	7011_vc0091_16k	女	专业娴熟/元气活力/ 激情饱满	广告营销/电商直播	乐观向上小姐姐

	姐姐			激情/元气/两		姐姐
68	琅琅上口女主播	7011_vc0062_16k	女	亲和力强/元气活力/激情饱满	广告营销/电商直播	琅琅上口女主播
69	铿锵有力气质女	7011_vc0093_16k	女	专业娴熟/元气活力/激情饱满	新闻主持/广告营销/教师讲座/电商直播	铿锵有力气质女
70	可爱邻家小妹	7011_vc0077_16k	女	亲和力强/幽默有趣/沉稳冷静	广告营销/电商直播	可爱邻家小妹
71	精力充沛妹子	7011_vc0030_16k	女	专业娴熟/元气活力/沉稳冷静	客服助理/新闻主持/广告营销/教师讲座/电商直播	精力充沛妹子
72	娇软可爱小甜妹	7011_vc0060_16k	女	亲和力强/幽默有趣/沉稳冷静	广告营销/电商直播	娇软可爱小甜妹
73	娇翘可爱小妹儿	7011_vc0129_16k	女	亲和力强/幽默有趣/沉稳冷静	广告营销/电商直播	娇翘可爱小妹儿
74	娇憨可爱萌妹	7011_vc0092_16k	女	亲和力强/元气活力/幽默有趣	广告营销/电商直播	娇憨可爱萌妹
75	坚毅果敢酷女孩	7011_vc0006_16k	女	元气活力/权威靠谱/激情饱满	广告营销/电商直播	坚毅果敢酷女孩
76	激进热情女声	7011_vc0025_16k	女	亲和力强/幽默有趣/激情饱满	客服助理/广告营销/电商直播	激进热情女声
77	活力满满女主播	7011_vc0072_16k	女	专业娴熟/幽默有趣/激情饱满	广告营销/电商直播	活力满满女主播
78	古灵精怪蜜友	7011_vc0094_16k	女	亲和力强/元气活力/幽默有趣	广告营销/电商直播	古灵精怪蜜友
79	端庄自信女主播	7011_vc0059_16k	女	亲和力强/元气活力/激情饱满	广告营销/电商直播	端庄自信女主播
80	东北调皮假小子	7011_vc0066_16k	女	亲和力强/元气活力/激情饱满	新闻主持/广告营销/教师讲座/电商直播	东北调皮假小子
81	东北大气女生	7011_vc0015_16k	女	专业娴熟/亲和力强/幽默有趣	广告营销/电商直播	东北大气女生
82	淡然豁达女老板	7011_vc0109_16k	女	专业娴熟/亲和力强/幽默有趣	新闻主持/广告营销/教师讲座/电商直播	淡然豁达女老板
83	大方利落大姐姐	7011_vc0050_16k	女	亲和力强/幽默有趣/沉稳冷静	广告营销/电商直播	大方利落大姐姐
84	大大咧咧妹子	7011_vc0012_16k	女	专业娴熟/元气活力/激情饱满	广告营销/电商直播	大大咧咧妹子
85	成熟知性女白领	7011_vc0081_16k	女	专业娴熟/权威靠谱/激情饱满	广告营销/电商直播	成熟知性女白领
86	成熟洒脱大姐姐	7011_vc0011_16k	女	专业娴熟/权威靠谱/激情饱满	广告营销/电商直播	成熟洒脱大姐姐
87	沉稳冷静女主播	7011_vc0014_16k	女	亲和力强/元气活力/沉稳冷静	客服助理/广告营销/电商直播	沉稳冷静女主播
88	沉稳果断女神	7011_vc0124_16k	女	亲和力强/元气活力/幽默有趣	广告营销/电商直播	沉稳果断女神
89	沉稳淡定酷妹	7011_vc0106_16k	女	专业娴熟/亲和力强/权威靠谱	广告营销/电商直播	沉稳淡定酷妹

90	才学过人俏佳人	7011_vc0128_16k	女	专业娴熟/幽默有趣/激情饱满	广告营销/电商直播	才学过人俏佳人
91	北京口音女声	7011_vc0100_16k	女	亲和力强/权威靠谱/沉稳冷静	广告营销/电商直播	北京口音女声
92	饱满有力活泼音	7011_vc0148_16k	女	亲和力强/元气活力/沉稳冷静	广告营销/电商直播	饱满有力活泼音
93	温柔质感女博主	7011_zhiningkoubo_16k	女	亲和力强/沉稳冷静/知性大方	广告营销/教师讲座	温柔质感女博主

男声音色

序号	音色名称	音色ID	性别	风格	适用场景	试听音频文件
1	度泽言-开朗	CAP_4193	男	温柔青年	客服助手	度泽言-开朗
2	度怀安	CAP_4195	男	磁性深情	客服助手	度怀安
3	度泽言-温暖	CAP_4179	男	温柔青年	客服助手	度泽言-温暖
4	度小新	4140	男	元气活力	广告营销	度小新
5	度星河	5135	男	沉稳冷静	客服助理/新闻主持/广告营销/教师讲座	度星河
6	度小凯	4123	男	激情饱满	客服助理/新闻主持/广告营销/教师讲座	度小凯
7	度逍遥	4003	男	权威靠谱/专业娴熟	客服助理/新闻主持/广告营销/教师讲座	度逍遥
8	度小彦	4129	男	元气活力	客服助理/广告营销	度小彦
9	度小贤	4115	男	权威靠谱/沉稳冷静	客服助理/新闻主持/广告营销/教师讲座	度小贤
10	度博文	4106	男	沉稳冷静	客服助理/新闻主持/广告营销/教师讲座	度博文
11	度清风	4143	男	元气活力	广告营销	度清风
12	度小科	4001	男	权威靠谱	广告营销	度小科
13	度小谋	4128	男	权威靠谱	广告营销	度小谋
14	度小笙	4127	男	沉稳冷静	客服助理/新闻主持/广告营销/教师讲座	度小笙
15	专业自信男主播	7011_moxingchuyi_16k	男	专业娴熟/亲和力强/沉稳冷静	广告营销/电商直播	专业自信男主播
16	自信专业男主播	7011_vc0135_16k	男	亲和力强/元气活力/激情饱满	客服助理/新闻主持/广告营销/教师讲座/电商直播	自信专业男主播
17	自信坦诚大男孩	7011_vc0104_16k	男	专业娴熟/元气活力/幽默有趣	广告营销/电商直播	自信坦诚大男孩
18	专业自信男青年	7011_vc0038_16k	男	亲和力强/幽默有趣/权威靠谱	客服助理/新闻主持/广告营销/教师讲座/电商直播	专业自信男青年
19	专业优雅大哥	7011_vc0058_16k	男	专业娴熟/权威靠谱/沉稳冷静	客服助理/广告营销/电商直播	专业优雅大哥

	男 男	女		亲和力/元气活力/幽默有趣		男 男
20	直接果断男 主播	7011_vc0041_16 k	男	亲和力强/元气活力/ 幽默有趣	广告营销/电商直播	直接果断男 主播
21	真诚自信男 青年	7011_vc0055_16 k	男	专业娴熟/元气活力/ 幽默有趣	广告营销/电商直播	真诚自信男 青年
22	友善直率大 男孩	7011_vc0132_16 k	男	亲和力强/幽默有趣/ 沉稳冷静	广告营销/电商直播	友善直率大 男孩
23	友善谦逊男 青年	7011_vc0085_16 k	男	专业娴熟/权威靠谱/ 激情饱满	客服助理/广告营销/电商直播	友善谦逊男 青年
24	硬朗自信小 哥哥	7011_vc0049_16 k	男	元气活力/幽默有趣/ 激情饱满	广告营销/电商直播	硬朗自信小 哥哥
25	阳光友善大 哥哥	7011_vc0123_16 k	男	专业娴熟/权威靠谱/ 沉稳冷静	广告营销/电商直播	阳光友善大 哥哥
26	阳光明朗少 年音	7011_vc0084_16 k	男	专业娴熟/权威靠谱/ 沉稳冷静	广告营销/电商直播	阳光明朗少 年音
27	阳光开朗大 哥哥	7011_vc0028_16 k	男	专业娴熟/权威靠谱/ 沉稳冷静	客服助理/新闻主持/广告营销/教师讲 座/电商直播	阳光开朗大 哥哥
28	雄浑宽广男 主播	7011_vc0147_16 k	男	专业娴熟/权威靠谱/ 沉稳冷静	新闻主持/广告营销/教师讲座/电商直 播	雄浑宽广男 主播
29	稳重成熟男 主播	7011_vc0054_16 k	男	专业娴熟/亲和力强/ 激情饱满	客服助理/新闻主持/广告营销/教师讲 座/电商直播	稳重成熟男 主播
30	温柔阳光男 闺蜜	7011_vc0107_16 k	男	元气活力/沉稳冷静/ 激情饱满	广告营销/电商直播	温柔阳光男 闺蜜
31	温柔体贴暖 男	7011_vc0061_16 k	男	专业娴熟/亲和力强/ 权威靠谱	客服助理/广告营销/电商直播	温柔体贴暖 男
32	温柔耐心大 男孩	7011_vc0113_16 k	男	专业娴熟/权威靠谱/ 沉稳冷静	广告营销/电商直播	温柔耐心大 男孩
33	温柔坚定男 声	7011_vc0064_16 k	男	专业娴熟/权威靠谱/ 激情饱满	广告营销/电商直播	温柔坚定男 声
34	温和谦虚小 哥哥	7011_vc0042_16 k	男	元气活力/权威靠谱/ 激情饱满	广告营销/电商直播	温和谦虚小 哥哥
35	头头是道讲 解员	7011_vc0079_16 k	男	亲和力强/元气活力/ 激情饱满	客服助理/新闻主持/广告营销/教师讲 座/电商直播	头头是道讲 解员
36	铁憨憨东北 大哥	7011_vc0018_16 k	男	亲和力强/权威靠谱/ 沉稳冷静	广告营销/电商直播	铁憨憨东北 大哥
37	谈笑风生大 哥哥	7011_vc0111_16 k	男	专业娴熟/权威靠谱/ 亲和力强	新闻主持/广告营销/教师讲座/电商直 播	谈笑风生大 哥哥
38	善解人意小 哥哥	7011_vc0016_16 k	男	专业娴熟/幽默有趣/ 激情饱满	广告营销/电商直播	善解人意小 哥哥
39	热情直率大 男孩	7011_vc0133_16 k	男	专业娴熟/亲和力强/ 激情饱满	广告营销/电商直播	热情直率大 男孩
40	热情洋溢男 主播	7011_vc0044_16 k	男	专业娴熟/幽默有趣/ 激情饱满	广告营销/电商直播	热情洋溢男 主播
41	清新阳光大 哥哥	7011_vc0083_16 k	男	亲和力强/幽默有趣/ 沉稳冷静	广告营销/电商直播	清新阳光大 哥哥

42	清新洒脱小哥哥	7011_vc0037_16k	男	亲和力强/元气活力/激情饱满	客服助理/广告营销/电商直播	清新洒脱小哥哥
43	清亮动人大男孩	7011_vc0145_16k	男	亲和力强/元气活力/权威靠谱	新闻主持/广告营销/教师讲座/电商直播	清亮动人大男孩
44	冷静认真男青年	7011_vc0082_16k	男	专业娴熟/亲和力强/沉稳冷静	新闻主持/广告营销/教师讲座/电商直播	冷静认真男青年
45	冷静豁达男主播	7011_vc0122_16k	男	元气活力/权威靠谱/激情饱满	广告营销/电商直播	冷静豁达男主播
46	乐观阳光小哥哥	7011_vc0089_16k	男	专业娴熟/亲和力强/沉稳冷静	广告营销/电商直播	乐观阳光小哥哥
47	侃侃而谈男主播	7011_vc0047_16k	男	元气活力/幽默有趣/权威靠谱	新闻主持/广告营销/教师讲座/电商直播	侃侃而谈男主播
48	开朗外向男闺蜜	7011_vc0036_16k	男	元气活力/幽默有趣/激情饱满	新闻主持/广告营销/教师讲座/电商直播	开朗外向男闺蜜
49	开朗京腔男主播	7011_vc0078_16k	男	专业娴熟/元气活力/权威靠谱	广告营销/电商直播	开朗京腔男主播
50	活力热情大哥哥	7011_vc0088_16k	男	元气活力/权威靠谱/激情饱满	广告营销/电商直播	活力热情大哥哥
51	浑厚有力小哥哥	7011_vc0144_16k	男	亲和力强/权威靠谱/激情饱满	新闻主持/广告营销/教师讲座/电商直播	浑厚有力小哥哥
52	东北磁性男声	7011_vc0067_16k	男	专业娴熟/权威靠谱/沉稳冷静	广告营销/电商直播	东北磁性男声
53	淡定温柔大哥哥	7011_vc0099_16k	男	亲和力强/幽默有趣/沉稳冷静	广告营销/电商直播	淡定温柔大哥哥
54	单纯憨厚小伙	7011_vc0101_16k	男	专业娴熟/元气活力/权威靠谱	广告营销/电商直播	单纯憨厚小伙
55	纯净清亮男主播	7011_vc0131_16k	男	亲和力强/权威靠谱/沉稳冷静	广告营销/电商直播	纯净清亮男主播
56	成熟谦逊大哥哥	7011_vc0134_16k	男	专业娴熟/权威靠谱/沉稳冷静	新闻主持/广告营销/教师讲座/电商直播	成熟谦逊大哥哥
57	沉着冷静讲解员	7011_vc0010_16k	男	元气活力/权威靠谱/沉稳冷静	广告营销/电商直播	沉着冷静讲解员
58	沉稳专业男主播	7011_vc0087_16k	男	专业娴熟/亲和力强/沉稳冷静	广告营销/电商直播	沉稳专业男主播
59	沉稳冷静大哥哥	7011_vc0043_16k	男	元气活力/权威靠谱/沉稳冷静	客服助理/新闻主持/广告营销/教师讲座/电商直播	沉稳冷静大哥哥
60	沉稳大气男老板	7011_vc0046_16k	男	专业娴熟/权威靠谱/沉稳冷静	广告营销/电商直播	沉稳大气男老板
61	自信沉稳精英男	7011_hanmiao_16k	男	专业娴熟/权威靠谱/沉稳冷静	新闻主持/广告营销/教师讲座	自信沉稳精英男

组件通用服务

接口通用说明

1. 接口鉴权

使用应用的 AppID、AppKey 和过期时间生成鉴权参数，在 Http 请求的 Header 传递鉴权参数进行鉴权。AppID 和 AppKey 的获取查看：[平台操作指引](#)

1.1 鉴权参数生成

鉴权参数格式：

```
AppId/Signature/ExpireTime
```

- AppId: 应用标识
- Signature: 使用 AppKey(应用密钥) 对 AppId 和 ExpireTime(过期时间) 进行 hmac_sha256 (<https://www.ietf.org/rfc/rfc2104.txt>) 运算生成的签名信息
- ExpireTime: 签名的过期时间
 - 格式：使用ISO8601格式表示(ISO_OFFSET_DATE_TIME)，具体的超时时间形如：2011-12-03T10:15:30+01:00
 - 根据业务场景设置，避免设置长期有效，以减小签名泄露的风险和影响

1.2 鉴权参数使用

需要加上以下的 HTTP 请求头，例如：

```
Authorization:
s-pgwi73tzrtwt/e9649e34ed495c753bdf214c81d67d1166f26e66776d51482aa10f1c70bc65e2/2023-07-07T08:03:10.
Z
```

1.3 示例代码

(1) 鉴权参数生成工具在线地址

<https://open.xiling.baidu.com/token-gen-tool.html>

(2) Postman Pre-script

```
var appId = " "; //填真实的appid
var appKey = " "; //填真实的appkey
var time = new Date(new Date().getTime() + 60 * 60 * 1000).toISOString();

const algorithm = CryptoJS.algo.SHA256;
const hmac = CryptoJS.algo.HMAC.create(algorithm, appKey);
hmac.update(appId + time);
var signature = hmac.finalize().toString();

pm.request.headers.add({
  key: 'Authorization',
  value: appId + "/" + signature + "/" + time
})
```

(3) Java demo

引入 commons-codec 包，版本：1.12：

Maven:

```
<dependency>
  <groupId>commons-codec</groupId>
  <artifactId>commons-codec</artifactId>
  <version>1.12</version>
</dependency>
```

Gradle:

```
implementation 'commons-codec:commons-codec:1.12'
```

```
package com.baidu.acg.plat.digitalhuman.agent.demonstration.service;
import org.apache.commons.codec.digest.HmacAlgorithms;
import org.apache.commons.codec.digest.HmacUtils;
import java.time.ZonedDateTime;
import java.time.format.DateTimeFormatter;
public class TokenDemo {
    public static void main(String[] args) throws Exception {
        String appld = " "; //填真实的appid
        String appKey = " "; //填真实的appkey
        // expiredTime eg: 2020-10-28T19:40:58.963441+08:00
        String expiredTime = ZonedDateTime.now().plusHours(1).format(DateTimeFormatter.ISO_OFFSET_DATE_TIME);
        // token eg: i-
        khpg99yk2j3gk/7dcefadebec0ee51f312ef1344d1d1cda9db4e35de7279ac7438823b5db96887/2020-10-
        28T19:40:58.963441+08:00
        String authorization = appld + "/" + hmacSha256(appKey, appld + expiredTime) + "/" + expiredTime;
    }
    private static String hmacSha256(String key, String data) {
        HmacUtils hmac = new HmacUtils(HmacAlgorithms.HMAC_SHA_256, key);
        return hmac.hmacHex(data);
    }
}
```

(4) Python demo

```
import hmac
import hashlib
import datetime

def hmac_sha256(key, data):
    return hmac.new(key.encode(), data.encode(), hashlib.sha256).hexdigest()

def main():
    app_id = " " # 填真实的appid
    app_key = "" # 填真实的appkey

    # expiredTime eg: 2020-10-28T19:40:58.963441+08:00
    expired_time = (datetime.datetime.now(datetime.timezone.utc) + datetime.timedelta(hours=1)).isoformat()

    # token eg: i-khpg99yk2j3gk/7dcefadebec0ee51f312ef1344d1d1cda9db4e35de7279ac7438823b5db96887/2020-
    10-28T19:40:58.963441+08:00
    authorization = f"{app_id}/{hmac_sha256(app_key, app_id + expired_time)}/{expired_time}"

    print(authorization)

if __name__ == "__main__":
    main()
```

2. 通用信息

2.1 通用请求字段

参数名	类型	必填	描述
requestId	string	否	用于标识这次请求，长度不超过 50。用于问题排查，如果不填系统会随机生成并在响应中返回。

下文若无特殊说明，接口请求参数描述会省略通用字段。

2.2 通用响应字段

参数名	类型	描述
requestId	string	请求标识
code	int	状态码，正常为 0
success	boolean	是否正常返回
message	object	提示信息
-- global	string	正常返回时为 "success"，异常时返回错误信息
result	object	响应数据，泛型类型，可能为 null

若无特殊说明，接口响应参数描述会省略这些通用字段，只针对 result 给出参数说明。

对于分页查询请求，接口响应的 result 字段会被 page 字段代替，说明如下：

参数名	类型	描述
page	object	
-- pageNo	int	页码，从 1 开始
-- pageSize	int	每页数量
-- totalCount	long	数据总数
-- result	object []	当前页数据列表，泛型类型

2.3 通用错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
14001	内部服务异常，请稍后再试
14002	内部服务异常，网络拥堵

🔗 2.4 任务回调接口

平台异步任务支持将任务完成事件发送到调用方提交任务填写的回调地址。回调消息只会发送一次，请求失败不会补发。POST到调用方提交任务时指定的 URL，长度不超过 1000。

为了提高回调请求的安全性，请求头会携带一个token，用户可以验证token的合法性决来判断该回调请求是否合法，token安全验证不是必须的操作，用户可以自主选择是否验证。

请求头 (Header) | 参数名 | 类型 | 是否必填 | 描述 | | token |
string | 是 | 请求的token，生成方式和本文档1.1章节中描述的鉴权生成规则一致，生成签名的参数时要对appkey和appid加 (callback) 的后缀，具体示例如下。用户接到请求后可以对token的合法性进行验证 |

(1) Postman Pre-script

```
var appId = " "; //填真实的appid
var appKey = " "; //填真实的appkey
var time = new Date(new Date().getTime() + 60 * 60 * 1000).toISOString();

const algorithm = CryptoJS.algo.SHA256;
const hmac = CryptoJS.algo.HMAC.create(algorithm, appKey+"_callback");
hmac.update(appId + "_callback_" + time);
var signature = hmac.finalize().toString();

pm.request.headers.add({
  key: 'Authorization',
  value: appId + "/" + signature + "/" + time
})
```

(2) Java demo

引入 commons-codec 包，版本：1.12：

Maven:

```
<dependency>
  <groupId>commons-codec</groupId>
  <artifactId>commons-codec</artifactId>
  <version>1.12</version>
</dependency>
```

Gradle:

```
implementation 'commons-codec:commons-codec:1.12'
```

```
package com.baidu.acg.plat.digitalhuman.agent.demonstration.service;
import org.apache.commons.codec.digest.HmacAlgorithms;
import org.apache.commons.codec.digest.HmacUtils;
import java.time.ZonedDateTime;
import java.time.format.DateTimeFormatter;
public class TokenDemo {
    public static void main(String[] args) throws Exception {
        String appId = " "; //填真实的appid
        String appKey = " "; //填真实的appkey
        // expiredTime eg: 2020-10-28T19:40:58.963441+08:00
        String expiredTime = ZonedDateTime.now().plusHours(1).format(DateTimeFormatter.ISO_OFFSET_DATE_TIME);
        // token eg: i-
        khpg99yk2j3gk/7dcefadebec0ee51f312ef1344d1d1cda9db4e35de7279ac7438823b5db96887/2020-10-
        28T19:40:58.963441+08:00
        String authorization = appId + "/" + hmacSha256(appKey + "_callback", appId + "_callback_" + expiredTime) + "/"
+ expiredTime;
    }
    private static String hmacSha256(String key, String data) {
        HmacUtils hmac = new HmacUtils(HmacAlgorithms.HMAC_SHA_256, key);
        return hmac.hmacHex(data);
    }
}
```

(3) Python demo

```
import hmac
import hashlib
import datetime

def hmac_sha256(key, data):
    return hmac.new(key.encode(), data.encode(), hashlib.sha256).hexdigest()

def main():
    app_id = " " # 填真实的appid
    app_key = "" # 填真实的appkey

    # expiredTime eg: 2020-10-28T19:40:58.963441+08:00
    expired_time = (datetime.datetime.now(datetime.timezone.utc) + datetime.timedelta(hours=1)).isoformat()

    # token eg: i-khpg99yk2j3gk/7dcefadebec0ee51f312ef1344d1d1cda9db4e35de7279ac7438823b5db96887/2020-
    10-28T19:40:58.963441+08:00
    authorization = f"{app_id}/{hmac_sha256(appKey + "_callback", appId + "_callback_" + expiredTime)}/{expired_time}"

    print(authorization)

if __name__ == "__main__":
    main()
```

请求参数列表（Body）

参数名	类型	是否必填	描述
requestId	string	是	请求唯一标识，用于问题排查
type	string	是	回调业务类型，比如单图视频合成任务，由具体业务接口定义
data	object	是	回调业务数据，由具体业务接口定义

SSML使用说明

语音合成（TTS）时一般输入纯文本即可，引擎内部会进行文本分析，预测合理的注音和韵律等。当需要能够更精确地控制合成效果时，比如指定人名中多音字读法、指定停顿时长等，可以使用对应 SSML（Speech Synthesis Markup Language，语音合成标记语言）标签来实现。

注意：SSML标签仅对中文生效，请勿在其他语言中使用，避免引发读音异常的问题。

这是一个 SSML 示例，所有文本包在 speak 顶层标签中，speak 标签可以不写，系统会补齐。

```
<speak>这是一个<say-as interpret-as="as-character">SSML</say-as>示例，我可以念多音字，<phoneme py="shan4">单</phoneme>老师说，<phoneme py="chan2">单</phoneme>于不会骑<phoneme py="dan1">单</phoneme>车。</speak>
```

支持的 SSML 标签：

分类	功能	示例	说明
多音字	多音字	<phoneme py='shan4'>单</phoneme>	通过拼音标注期望的发音
数字	整数	<say-as type="number:ordinal">123</say-as>	读作：一百二十三
数字	数字串	<say-as type="number:digits">123</say-as>	读作：一二三
数字	比分	<say-as type="number:score">12:123</say-as>	读作：十二比一百二十三
数字	分数	<say-as type="number:fraction">12/123</say-as>	读作：一百二十三分之十二
数字	电话号码	<say-as type="telephone">110</say-as>	读作：幺幺零
英文	字母	<say-as interpret-as="as-character">BAT</say-as>	
英文	单词	<say-as interpret-as="as-word">BAT</say-as>	
停顿	停顿	<silence time="0.5s"></silence>	最大值10秒

文件上传接口

接口说明

接口地址域名：<https://open.xiling.baidu.com>

用于上传形象克隆和声音克隆等接口需要的音视频文件，用户先调用上传接口上传文件，拿到 fileId 后在业务接口使用。

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

接口列表

文件上传

接口说明

上传文件，返回文件Id。

🔗 接口协议

POST /api/digitalhuman/open/v1/file/upload

🔗 请求参数

参数	类型	必填	说明
file	MultipartFile	是	文件内容，文件大小、格式等要求见业务接口的文档说明
providerType	String	是	上传文件的业务类型，请参考具体业务接口的说明填写。
sourceFileName	String	是	上传的文件名，必须填写正确的文件名称及后缀比如：test.mp4。

🔗 返回参数

参数	类型	必填	说明
fileId	String	是	文件ID，后续用于业务接口，文件默认保存 7 天，请及时使用
fileName	String	是	文件的原始名称，用户上传时的原文件名

🔗 请求示例

Postman 示例

SAAS / 开放平台 / 文件上传

Save

Send

POST (api-gateway)/api/digitalhuman/open/v1/file/upload?providerType=OPEN_CUSTOMIZATION_PICTURE&sourceFileName=1111.mp4

Params Authorization Headers (9) Body Pre-request Script Tests Settings Cookies

none form-data x-www-form-urlencoded raw binary GraphQL

Key	Value	Description	Bulk Edit
file	66504c62c65e703256eefaf0_427.mp4		
Key	Text Value	Description	

Body Cookies (1) Headers (5) Test Results

Status: 200 OK Time: 8.39 s Size: 352 B Save as example

Pretty Raw Preview Visualize JSON

```
1 {
2   "code": 0,
3   "message": {
4     "global": "success"
5   },
6   "result": {
7     "fileId": "65ecdcbc-639c-4a73-8ac3-0c3d45c456da",
8     "fileName": "1111.mp4"
9   },
10  "requestId": "6ad2b30f-bf4d-4bab-baa7-62142ecd7f21",
11  "success": true
12 }
```

Java 代码示例

```
OkHttpClient client = new OkHttpClient().newBuilder()
    .build();
RequestBody body = new MultipartBody.Builder().setType(MultipartBody.FORM)
    .addFormDataPart("file", "664ddd142a021b7ca3ff7e94_217.wav",
        RequestBody.create(MediaType.parse("audio/wav"),
            new File("/Users/xxx/Downloads/664ddd142a021b7ca3ff7e94_217.wav")))
    .build();
Request request = new Request.Builder()
    .url("https://open.xiling.baidu.com/api/digitalhuman/open/v1/file/upload?providerType"
        + "=OPEN_CUSTOMIZATION_PICTURE&sourceFileName=1111.mp4")
    .method("POST", body)
    .addHeader("Authorization", "your Authorization")
    .build();
Response response = client.newCall(request).execute();
CommonResponse<FileStorageResponse> fileStorageResponseCommonResponse =
    JsonUtil.readValue(response.body().string(), new TypeReference<CommonResponse<FileStorageResponse>>() {
    });
System.out.println(JsonUtil.writeValueAsStringIgnoreNull(fileStorageResponseCommonResponse));
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "fileId": "f2c1b49b-a62c-407b-95f8-5dd7085a44eb",
    "fileName": "664ddd142a021b7ca3ff7e94_217.wav"
  },
  "requestId": "b3c53f39-cb78-4bab-abbf-077ecc12e3cf",
  "success": true
}
```

错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
14001	内部服务异常，请稍后再试
14002	内部服务异常，网络拥堵
30001	用户上传文件解析异常
30002	获取文件宽高或时长异常
30003	上传文件类型异常
30004	上传文件大小超过最大限制
30005	上传文件时长超过最大限制
30006	上传文件像素数超过最大限制
30007	上传文件到bos异常

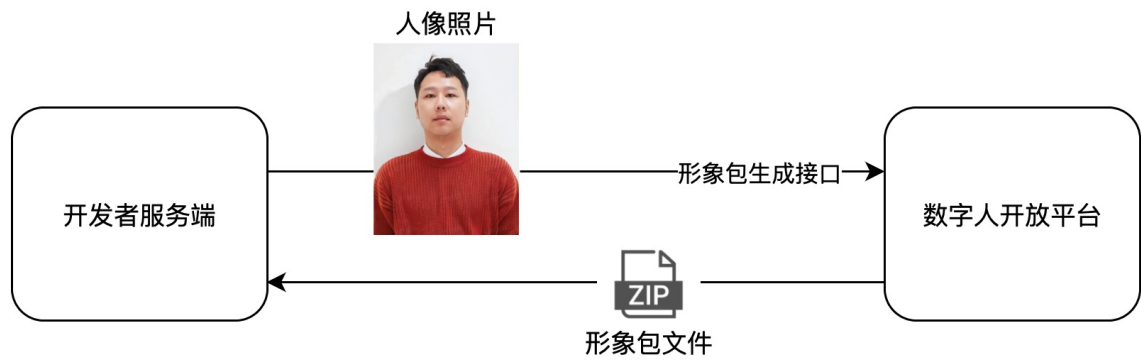
定制服务

照片数字人形象包定制接口

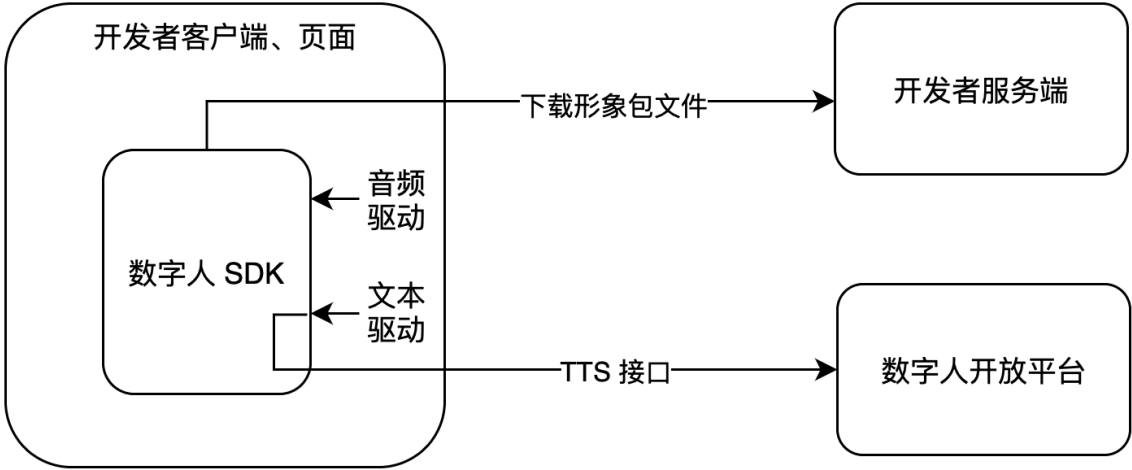
1 使用说明

1.1 使用流程

1、形象定制：**本文档为照片数字人形象定制接口文档**，服务端调用数字人开放平台“形象定制-照片数字人”接口，上传人像照片获取形象包文件。已生成形象包需要通过SDK驱动使用，不可单独使用。



2、SDK调用：客户端或页面集成[端渲染交互组件-照片数字人SDK](#)加载形象包，本地驱动渲染数字人。支持音频和文本驱动。音频可以使用真人音频或自行接入的第三方TTS的音频。文本驱动时 SDK 内部会调用数字人开放平台 TTS 服务驱动数字人。



1.2 照片要求

注意：输入的照片需符合以下规范，否则形象包渲染效果可能异常

- 图片格式：png，jpg
- 分辨率：分辨率不超过1080*1920px（建议人脸区域不超过256*256px）
- 照片要求：
 - 1、照片清晰，光线自然。
 - 2、人脸无遮挡、五官无遮挡。
 - 3、直面镜头，人脸各方向转角控制在15°以内，不能出现侧脸、歪头、低头、仰头等情况。
 - 4、不露牙齿，人脸表情尽量放松，避免抿嘴、张嘴等情况，允许小幅度微笑。
 - 5、避免人物头部穿戴复杂头饰、帽子。
 - 6、避免出现丸子头、马尾等高颅顶发型。
 - 7、避免人物头发过于蓬松。
 - 8、避免使用动漫人物等单线条唇形，需保证人物图片具备正常的真人唇型，否则无法保证效果。



1.3 购买入口：

组件商店：<https://xiling.cloud.baidu.com/open/widgetStore/list>

2 接口说明

接口调用地址：<https://open.xiling.baidu.com> 使用图片训练人像包

1. 调用人像包训练接口，提交图片，返回任务ID
2. 调用人像包查询接口，提交任务ID，获取训练信息，包含状态及人像包地址等

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

3 接口列表

3.1 提交人像训练接口

3.1.1 接口说明

调用人像包训练接口，提交图片，返回任务ID

3.1.2 接口协议

POST /api/digitalhuman/open/v1/figure/image/create

Header Content-Type: application/json;charset=utf-8

3.1.3 请求参数

参数	类型	必填	说明
figureImageURL	string	是	人像图片 url，url 字符长度不超过 1000，图片需要包含清晰的人脸，图片格式要求：文件大小不超过 1.5 MB；支持的图片格式：png、jpg、jpeg；当图片是包含透明通道的 png 格式，会返回 webm 格式透明背景视频，其他情况返回 mp4 格式视频。
callbackUrl	string	否	接口调用方接受任务回调通知的url，规则和接口说明参考

3.1.4 回调参数

- type: PICTURE_FIGURE_ADD
- data:
-

参数名	类型	是否必填	描述
taskId	string	是	任务ID
status	string	是	状态：
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息 1. SUCCESS（合成成功） 2. FAILED（合成失败）
figureUrl	string	否	形象包文件地址，文件会保存 7 天

3.1.5 返回参数

参数	类型	必填	说明
taskId	string	是	任务ID，用于调用任务查询接口查询任务状态

3.1.6 请求示例

```
{
  "requestId": "xxx11",
  "figureImageUrl": "https://meta-human-editor-prd.cdn.bcebos.com/ca1d70b6f73d44bebd2aba1f3deb60bb/draftThumbnail/dt-qfdkzu6nhawzayqb5.png",
  "callbackUrl": "http://api-gateway:8080/api/digitalhuman/internal/api-gateway/v1/callback/vp/test"
}
```

3.1.7 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "fig-qgti819vzahqxp3",
  },
  "requestId": "xxx11",
  "success": true
}
```

3.1.8 错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
14001	内部服务异常，请稍后再试
14002	内部服务异常，网络拥堵
20001	未检查到人脸
20034	图片下载异常
21001	形象包任务提交失败

3.2 查询任务

3.2.1 接口说明

使用任务ID查询任务状态

3.2.2 接口协议

GET /api/digitalhuman/open/v1/figure/image/task

3.2.3 GET 请求参数

参数	类型	必填	说明
taskId	string	是	任务ID

3.2.4 返回参数

参数	类型	必填	说明
taskId	string	是	任务ID
status	string	是	状态：
failedCode	int	否	失败错误码 1. SUBMIT（已提交待合成） 2. GENERATING（合成中） 3. SUCCESS（合成成功） 4. FAILED（合成失败）
failedMessage	string	否	失败错误信息
figureUrl	string	否	形象包文件地址，文件会保存 7 天
createTime	string	是	创建时间，示例：2024-04-29T20:28:06
updateTime	string	是	更新时间，示例：2024-04-29T20:28:06

3.2.5 请求示例

GET /api/digitalhuman/open/v1/figure/image/task?taskId=abc123&requestId=123e4567-e89b-12d3-a456-426614174000

3.2.6 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "fig-qgti819vzahqxb3",
    "status": "GENERATING",
    "failedCode": 0,
    "failedMessage": "",
    "figureUrl": null,
    "createTime": "2024-07-18T08:17:56",
    "updateTime": "2024-07-18T08:17:56"
  },
  "requestId": "fdc66116-762d-4a82-a082-e83357a3167e",
  "success": true
}
```

3.3 提交照片抠图接口

3.3.1 接口说明

调用照片抠图接口，提交图片，返回抠图结果

3.3.2 接口协议

POST /api/digitalhuman/open/v1/photo/bodyseg

Header Content-Type: application/json;charset=utf-8

3.3.3 请求参数

参数	类型	必填	说明
imgUrl	string	是	人像图片文件url，必须为 BOS 文件地址（通过 https://cloud.baidu.com/product/bos.html 上传后获得） 图片需要包含清晰的人脸，图片格式要求： 1. 文件大小不超过 3 MB 2. 支持的图片格式：png、jpg、jpeg； 3. 图片分辨率不超过1080P 图片内容要求： 1. 照片清晰，只有单个人脸 2. 面部正对镜头，轮廓完整 3. 面部无明显阴影，无头发遮挡
customizeType	string	是	定制类型： OPEN_PHOTO_KEYING（照片抠像）

3.3.4 返回参数

参数	类型	必填	说明
imgResultUrl	string	是	抠图结果。文件保存 7 天，请及时使用

3.3.5 请求示例

```
{
  "requestId": "xxxxx1",
  "imgUrl": "xxxxxx",
  "customizeType": "OPEN_PHOTO_KEYING"
}
```

3.3.6 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "imgResultUrl": "https://meta-human-editor-prd.cdn.bcebos.com/ca1d70b6f73d44bebd2aba1f3deb60bb/draftThumbnail/dt-qfdkzu6nhawzayqb5.png"
  },
  "requestId": "xxxx11",
  "success": true
}
```

3.3.7 错误码

错误码	描述
10006	参数校验不通过
10011	没有购买对应的商品
25001	图片抠像失败
25003	图片格式不正确，请检查
25004	图片大小过大，请检查
25005	图片分辨率过大，请检查
25006	图片解析错误，请检查
25007	图片地址必须为BOS文件地址，请检查
25008	图片下载失败，请检查图片是否存在

2D小样本数字人（通用口型）拍摄指南

定制2D小样本数字人（通用口型）时，我们需要您提供一段10秒-4分钟自然闭嘴的真人视频。为了提升最终生成的数字人效果，我们特别准备了以下拍摄指南，供您参考：

示例视频部分片段：

- 实景拍摄：训练固定背景数字人（保留原背景）
- 绿幕拍摄：训练透明背景数字人（可替换任意背景）

1 拍摄准备

1. 模特服饰

- 可按照模特喜好选择，与背景颜色风格协调、适合实际应用场景即可
- 合身，不能过大或过紧
- 整洁大方，无明显褶皱
- 穿低领衣服，露出脖子
- 避免碎花、格纹、条纹等大面积花纹的衣服，避免摩尔纹产生
- 建议着深色服装，或者选择其他与背景颜色反差大的服饰颜色
- 不建议选择白色、反光材质、过于光滑面料，会反绿（绿幕拍摄需注意）
- 服饰中不存在绿色系服饰，不存在纱质材料，会漏绿（绿幕拍摄需注意）

2. 模特妆造

- 妆容自然大方，有亲和力，无明显突兀感
- 面部白皙且有立体感，但不宜有过重阴影
- 如果存在瑕疵，例如皮肤不平整，长痘，手部存在血丝，需要遮瑕
- 嘴唇过薄可通过口红补救
- 露出双耳，可佩戴固定耳饰，不建议佩戴易晃动耳饰
- 注意手部、脖子、腿等部位，需要保持与面部肤色一致
- 头发正面看无碎发、无毛刺、无镂空，饱满有弧度（绿幕拍摄需注意）

3. 布景：

- 实景拍摄：背景整洁、无密集纹理，适合实际应用场景
- 绿幕拍摄：绿幕颜色统一，且平整无褶皱；建议人物和绿幕距离大于2米，避免人物反绿

4. 灯光：

- 用灯光从斜前侧方照亮模特面部，使主播在画面中更突出
- 在灯光下，模特皮肤白皙透亮，避免脸部反光、肤色暗沉或曝光
- 模特光线均匀，尤其注意脸部，颈部，手部，腿部的光照一致性
- 背景光照均匀，并与人像轮廓形成清晰边界
- 灯光一旦确定，切勿再次调整，否则全部重拍

5. 设备：

- 建议拍摄的分辨率不低于1080P，关闭HDR模式，关闭log模式
- 比例：横屏为16:9，竖屏为9:16（若最终的场景用于竖屏，建议使用竖屏拍摄；横屏同理）
- 角度：镜头高度与主播眼睛保持同一水平线，模特在头部保持自然的情况下平视镜头，建议正对镜头，侧身不超过10度，不要仰头、低头、仰视、俯视
- 构图：将主播放置于画面中间，确保五官清晰、面部的无遮挡
- 可使用三脚架或稳定器固定拍摄设备，以减少画面抖动，确保视频流畅稳定

2 形象拍摄

1. 面部神情：头部保持相对稳定，模仿说话状态轻微晃动，面带微笑，面部肌肉放松，注意眼睛一直看镜头，避免张口露齿、抿嘴、眼睛四处看（无需说话）
 2. 模特动作：越自然越好，以下建议仅供参考
- 尽量选择通用的小幅动作，比如摊手、合手、上下叠手等（细微的手部动作会更自然），注意避免摆手、摇头、比数字等不通用动作，动作的含义需适配数字人播报的所有文本及场景，生成的数字人姿态动作将与视频完全一致，请确保拍摄时的姿态可用于后续内容创作
 - 手势动作避免遮挡嘴巴、面部、脖子周围，也避免出画面
 - 可根据应用场景决定动作
 - 视频场景：更适合把动作幅度控制在身前，增加一些自然的手部小动作、偶尔有摊手等较大幅度动作会更有聊天感
 - 交互场景：整个视频中穿插「聆听动作」和「手势动作」两种，一镜到底，总时长建议40s-2min。录制开始「聆听动作」停顿 10-20s，然后开始做动作，最后需要回到第一个「聆听动作」（位置动作完全一致）。
 - 聆听动作：指未讲话时的聆听状态，可以轻微点头、侧头，身体正常的呼吸轻晃，过程中归位的手部动作保证基本位置一致（胳膊、小臂、手腕位置一致），手部动作可变。每个5-6s。
 - 手势动作：讲话时的摆手状态，根据个人讲话习惯设计不同的手势，避免有指向意义的动作、避免连续长动作。每个2-4s。
 - 直播场景：可自然闭嘴模拟一场真实的直播，可让主播假装在介绍产品、与用户互动等

tips：

- 可以连续拍摄4分钟左右，为剪辑提供更多的素材
- 可以播放一段文本内容，主播想象自己在播报并自然配合神情和手势



3 后期处理

3.1 剪辑-调色美颜

- 需掐头去尾剪辑，最终提供的底板视频需全程无拼接，跳帧对视频效果影响较大
- 开头保留2-3秒没有手势的静默状态，可以使动作自然开始
- 片段删减说明
 - 不建议保留：不通用动作、指向性动作、否定动作、眼神乱瞟等不合时宜的动作
 - 可以保留：偶尔几秒钟的露齿对训练效果基本没影响
- 主播动作过快或过慢时，可以适当变速，目前来看0.8-1.3倍速基本不会影响效果

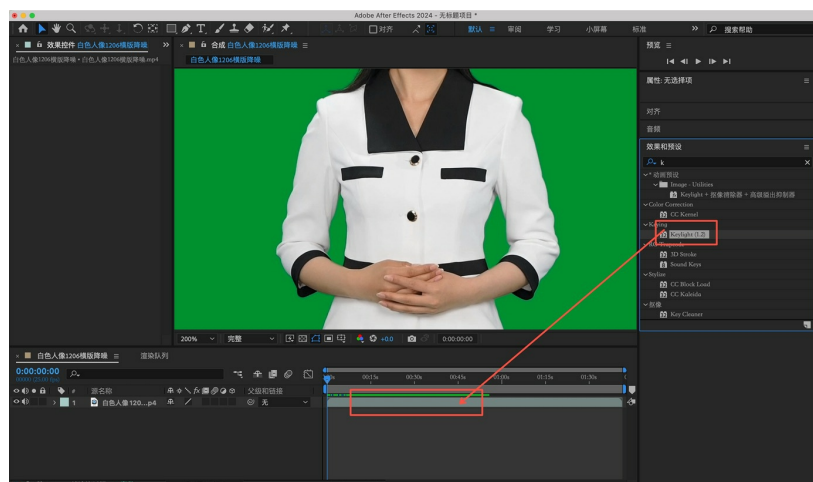
如果主播有美颜需求，也可以在剪辑时用美颜软件还原美貌，数字人训练效果会与视频素材保持一致

- 最终底板视频时长要求1-4分钟，2-3分钟为宜
- 若无需抠绿，可直接导出视频，实景拍摄的底板视频规格要求如本文【4.2 视频规格要求】

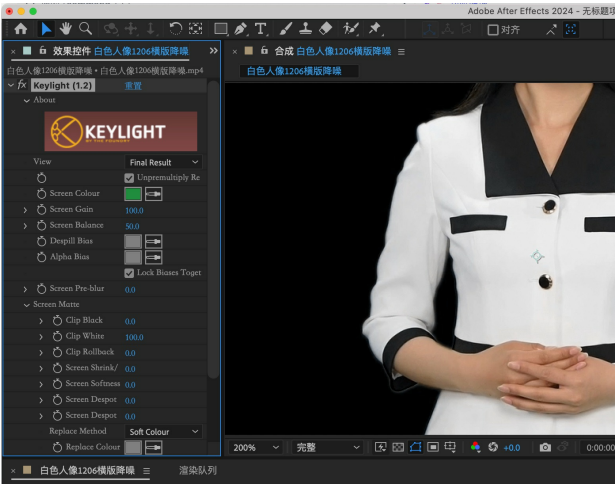
3.2 抠绿（绿幕拍摄）

以下为AE（Adobe After Effects）抠绿方式及导出视频规格，也可选择其他软件进行抠绿处理。

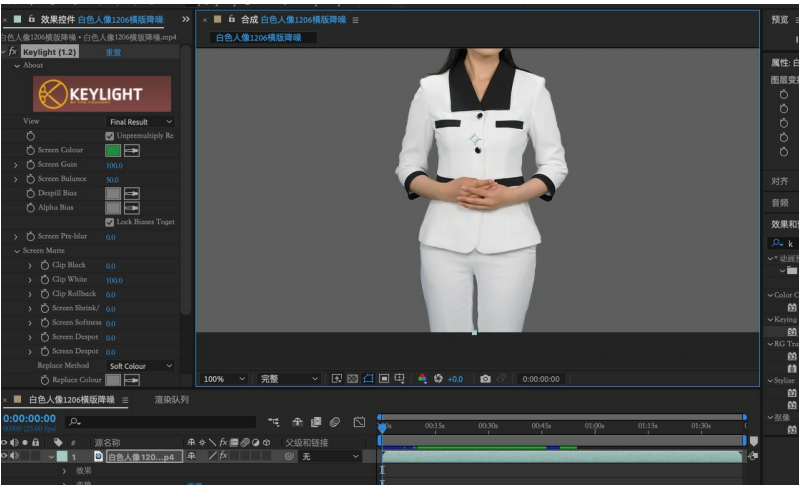
1. 对视频添加超级键效果：



2. 调整各种参数：抠除绿色背景，去除模特脸部以及身上绿色反光



3. 检查抠绿效果：抠绿后可换浅灰色背景，检查抠绿是否干净以及是否有该保留的部分被抠除



4. 导出视频：

方案一（推荐）：

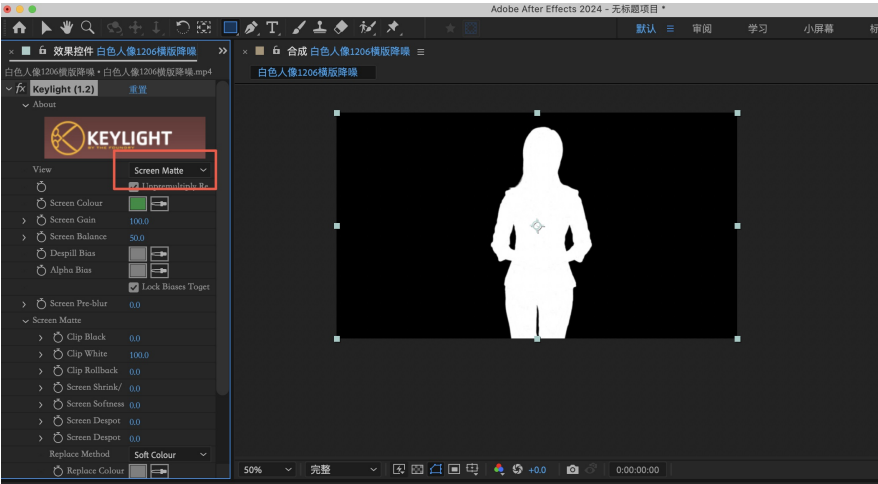
- 第一步：导出透明背景的带alpha通道（预乘模式）的视频
 - 格式：mov（Apple Prores 4444或Apple Prores 4444 XQ编码）
 - 帧率：25fps
 - 分辨率：1080P、2K、4K（2K、4K人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景）



- 第二步：使用ffmpeg将mov转换成带alpha通道的webm视频，webm视频规格要求如本文【4.2 视频规格要求】

```
ffmpeg -i input.mov -c:v libvpx-vp9 -pix_fmt yuva420p -lossless 1 -b:v 0 output.webm
```

方案二：导出浅灰色背景的底板视频和mask视频，选择【mask模式】方式如下，视频规格要求如本文【4.2 视频规格要求】



4 视频自检

在提交视频前，请务必逐条确认视频素材是否满足以下要求。

4.1 内容要求：

- 自然闭嘴，不说话、不张嘴、不露牙齿
- 动作自然，无指向性动作
- 头部无大角度转动，脸部全程无遮挡
- 提供的视频全程无拼接、无跳帧
- 人像边缘平滑自然，没有绿色残留或锯齿状边缘（若抠绿需注意）
- 穿低领衣服，露出脖子
- 光线充足，脸部整体光线均匀
- 保持妆造、光线、景别不变

4.2 视频规格要求：

	固定背景（保留原背景）	透明背景（可替换任意背景）	
拍摄布景	实景拍摄	绿幕拍摄	
底板视频规格	* 数量：1个 * 大小：3GB以内 * 时长：通用口型需1–4分钟 * 格式：mp4（H.264编码） * 分辨率：1080P、2K、4K（2K、4K人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景） * 比例：16：9或9：16 * 帧率：25 fps * 色彩标准：SDR–Rec.709（非HDR）	方案一： * 数量：1个 * 大小：3GB以内 * 时长：通用口型需1–4分钟 * 格式：webm（VP8、VP9编码） * 分辨率：1080P、2K、4K（2K、4K人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景） * 比例：16：9或9：16 * 帧率：25 fps * 色彩标准：SDR–Rec.709（非HDR）	方案二： * 数量：1个 * 大小：3GB以内 * 时长：通用口型需1–4分钟 * 格式：mp4（H.264编码） * 分辨率：1080P、2K、4K（2K、4K人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景） * 比例：16：9或9：16 * 帧率：25 fps * 色彩标准：SDR–Rec.709（非HDR）
mask视频规格	无需上传	无需上传	* 数量：1个 * 大小：3GB以内 * 时长、格式、分辨率、比例、帧率：与底板视频相同

2D小样本数字人（专属口型）拍摄指南

定制2D小样本数字人（专属口型）时，我们需要您提供一段10秒-4分钟的闭嘴动作底板视频，和一段4-8分钟的真人口播唇动视频，为了提升最终生成的数字人效果，我们特别准备了以下拍摄指南，供您参考：

示例视频部分片段：

绿幕拍摄：训练透明背景数字人（可替换任意背景）

- 底板视频：一段10秒-4分钟的闭嘴视频，用于人像动作训练，视频需保留音轨，训练出的数字人动作会与该底板完全一致
- 唇动视频：一段4-8分钟的真人口播视频，用于数字人口型训练，视频需保留音轨，音画同步，模特声音清晰无环境噪音

🔗 1 拍摄准备

1. 模特服饰

- 可按照模特喜好选择，与背景颜色风格协调、适合实际应用场景即可
- 合身，不能过大或过紧
- 整洁大方，无明显褶皱
- 穿低领衣服，露出脖子
- 避免碎花、格纹、条纹等大面积花纹的衣服，避免摩尔纹产生
- 建议着深色服装，或者选择其他与背景颜色反差大的服饰颜色
- 不建议选择白色、反光材质、过于光滑面料，避免金属配饰，会反绿（绿幕拍摄需注意）
- 服饰中不存在绿色系服饰，不存在纱质材料，会漏绿（绿幕拍摄需注意）

2. 模特妆造

- 无美颜情况下上镜效果好，眉毛无遮挡
- 妆容自然大方，有亲和力，无明显突兀感
- 面部白皙且有立体感，但不宜有过重阴影
- 如果存在瑕疵，例如皮肤不平整，长痘，手部存在血丝，需要遮瑕
- 嘴唇过薄可通过口红补救

- 露出双耳，可佩戴固定耳饰，不建议佩戴易晃动耳饰
- 演员应避免带眼镜，镜片极易反绿，影响后期抠图效果
- 注意手部、脖子、腿等部位，需要保持与面部肤色一致
- 头发正面看无碎发、无毛刺、无镂空，饱满有弧度（绿幕拍摄需注意）

3. 播报文本：

- 准备4-8分钟约1500字以上可读性较强的语料
- 文本内容最好贴合实际应用场景
- 建议模特在录制前熟读文本

4. 布景：

- 实景拍摄：背景整洁、无密集纹理，适合实际应用场景
- 绿幕拍摄：绿幕颜色统一，且平整无褶皱；建议人物和绿幕距离大于2米，避免人物反绿

5. 灯光：

- 用灯光从斜前侧方照亮模特面部，使主播在画面中更突出
- 在灯光下，模特皮肤白皙透亮，避免脸部反光、肤色暗沉或曝光
- 模特光线均匀，尤其注意脸部，颈部，手部，腿部的光照一致性
- 背景光照均匀，并与人像轮廓形成清晰边界
- 灯光一旦确定，切勿再次调整，否则全部重拍

6. 设备：

- 建议拍摄的分辨率不低于1080P，关闭HDR模式，关闭log模式
- 需准备提词器，放映预先准备的播报文本
- 比例：横屏为16:9，竖屏为9:16（若最终的场景用于竖屏，建议使用竖屏拍摄；横屏同理）
- 角度：镜头高度与主播眼睛保持同一水平线，模特在头部保持自然的情况下平视镜头，建议正对镜头，侧身不超过10度，不要仰头、低头、仰视、俯视
- 构图：将主播放置于画面中间，确保五官清晰、面部无遮挡
- 可使用三脚架或稳定器固定拍摄设备，以减少画面抖动，确保视频流畅稳定

2 形象拍摄

底板视频：

1. 面部神情：头部保持相对稳定，模仿说话状态轻微晃动，面带微笑，面部肌肉放松，注意眼睛一直看镜头，避免张口露齿、抿嘴、眼睛四处看（无需说话）
2. 模特动作：越自然越好，以下建议仅供参考
 - 尽量选择通用的小幅动作，比如摊手、合手、上下叠手等（细微的手部动作会更自然），注意避免摆手、摇头、比数字等不通用动作，动作的含义需适配数字人播报的所有文本及场景，生成的数字人姿态动作将与视频完全一致，请确保拍摄时的姿态可用于后续内容创作
 - 手势动作避免遮挡嘴巴、面部、脖子周围，也避免出画面

- 可根据应用场景决定动作
 - 视频场景：更适合把动作幅度控制在身前，增加一些自然的手部小动作、偶尔有摊手等较大幅度动作会更有聊天感
 - 交互场景：整个视频中穿插「聆听动作」和「手势动作」两种，一镜到底，总时长建议40s-2min。录制开始「聆听动作」停顿 10-20s，然后开始做动作，最后需要回到第一个「聆听动作」（位置动作完全一致）。
 - 聆听动作：指未讲话时的聆听状态，可以轻微点头、侧头，身体正常的呼吸轻晃，过程中归位的手部动作保证基本位置一致（胳膊、小臂、手腕位置一致），手部动作可变。每个5-6s。
 - 手势动作：讲话时的摆手状态，根据个人讲话习惯设计不同的手势，避免有指向意义的动作、避免连续长动作。每个2-4s。
 - 直播场景：可自然闭嘴模拟一场真实的直播，可让主播假装在介绍产品、与用户互动等

🔗 唇动视频：

1. 唇动视频模特形象、机位、灯光等需与底板视频完全一致，建议模特连续拍摄底板视频与唇动视频，后期剪辑
2. 牙齿整齐，说话嘴张开上下牙齿全部漏出，避免讲话不漏牙齿或者只漏上牙或下牙。在句与句之间多停顿，停顿时自然闭嘴

tips：

- 可以连续拍摄10分钟左右，为剪辑提供更多的素材

🔗 3 后期处理

🔗 3.1 剪辑-调色美颜

- 需掐头去尾剪辑，最终提供的底板视频需全程无拼接，跳帧对视频效果影响较大
- 开头保留2-3秒没有手势的静默状态，可以使动作自然开始
- 片段删减说明
 - 不建议保留：不通用动作、指向性动作、否定动作、眼神乱瞟、笑场等不合时宜的动作，有较大噪音、其他人说话等片段
 - 可以保留：录制过程中微小卡顿，偶尔念错对训练效果基本没影响

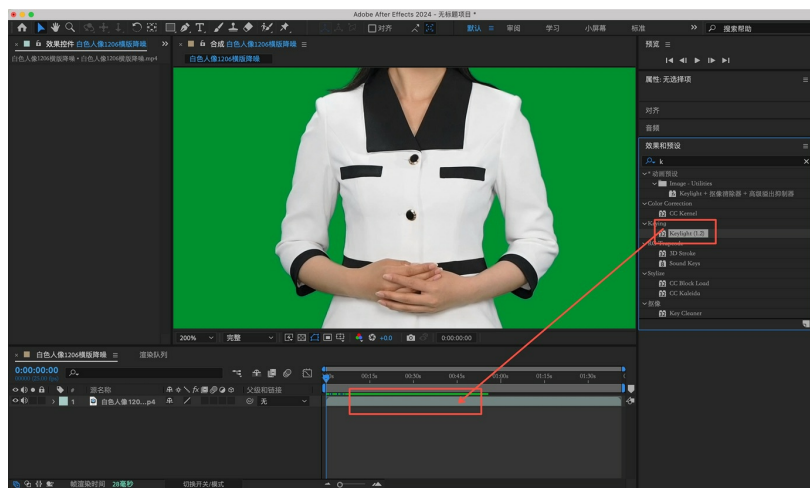
如果主播有美颜需求，也可以在剪辑时用美颜软件还原美貌，数字人训练效果会与视频素材保持一致

- 若无需抠绿，可直接导出视频，实景拍摄的底板视频规格要求如本文【4.2 视频规格要求】

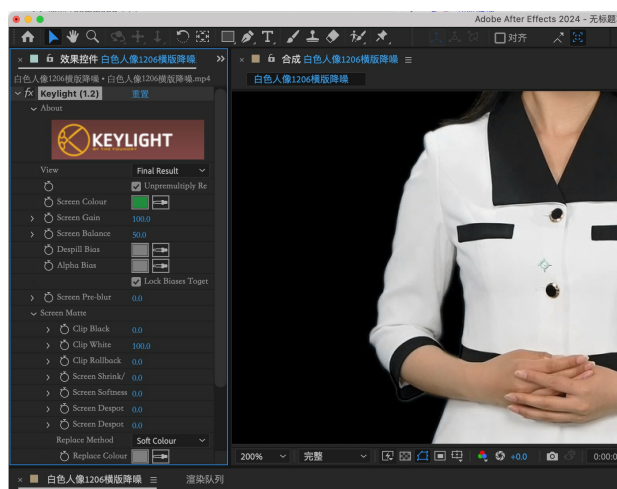
🔗 3.2 抠绿（绿幕拍摄）

以下为AE（Adobe After Effects）抠绿方式及导出视频规格，也可选择其他软件进行抠绿处理。

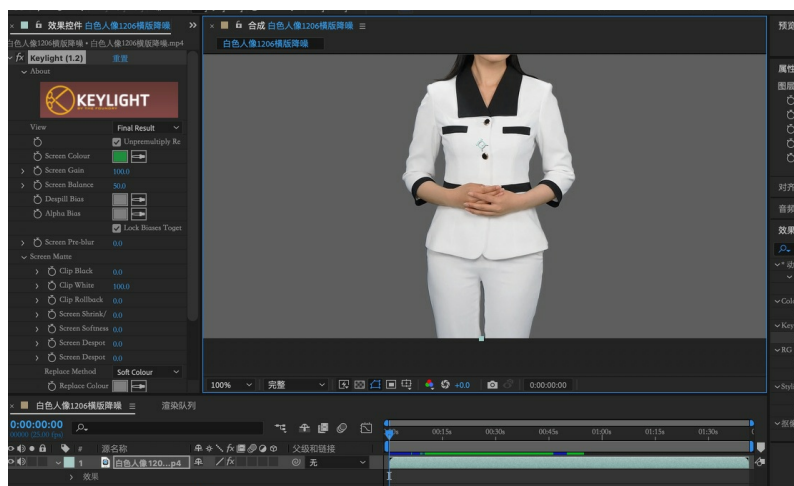
1. 对视频添加超级键效果：



2. 调整各种参数：抠除绿色背景，去除模特脸部以及身上绿色反光



3. 检查抠绿效果：抠绿后可换浅灰色背景，检查抠绿是否干净以及是否有该保留的部分被抠除



4. 导出视频：

方案一（推荐）：

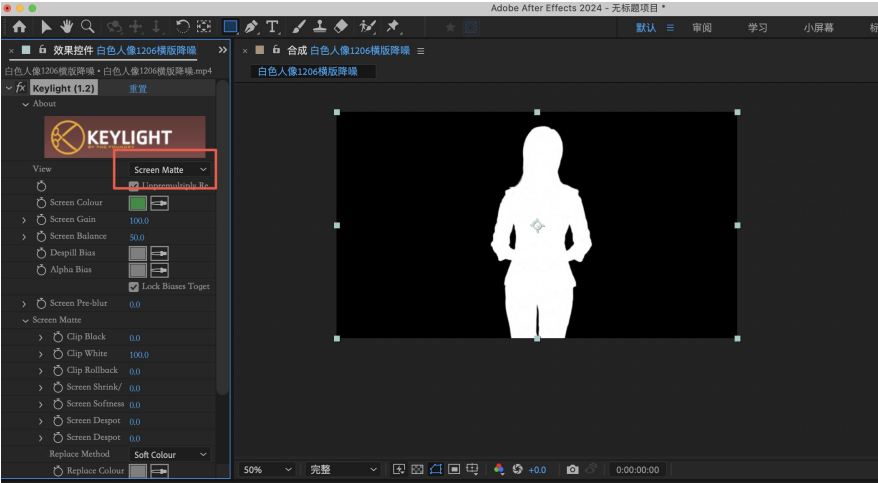
- 第一步：导出透明背景的带alpha通道（预乘模式）的视频
 - 格式：mov（Apple Prores 4444或Apple Prores 4444 XQ编码）
 - 帧率：25fps
 - 分辨率：1080P、2K、4K（2K、4K人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推理场景）



- 第二步：使用ffmpeg将mov转换成带alpha通道的webm视频，webm视频规格要求如本文【4.2 视频规格要求】

ffmpeg -i input.mov -c:v libvpx-vp9 -pix_fmt yuva420p -lossless 1 -b:v 0 output.webm

方案二：导出浅灰色背景的底板视频和mask视频，选择【mask模式】方式如下，视频规格要求如本文【4.2 视频规格要求】



4 视频自检

在提交视频前，请务必逐条确认视频素材是否满足以下要求。

4.1 内容要求：

- 动作自然，无指向性动作
- 在句与句之间多停顿，停顿时自然闭嘴
- 头部无大角度转动，脸部全程无遮挡
- 提供的视频全程无拼接、无跳帧
- 视频音画同步，无明显噪音
- 人像边缘平滑自然，没有绿色残留或锯齿状边缘（若抠绿需注意）
- 穿低领衣服，露出脖子
- 光线充足，脸部整体光线均匀
- 保持妆造、光线、景别不变

4.2 视频规格要求：

	固定背景（保留原背景）	透明背景（可替换任意背景）	
拍摄布景	实景拍摄	绿幕拍摄	
底板视频规格	* 数量：1个 * 大小：3GB以内 * 时长：10秒 – 4分钟闭嘴视频 * 格式：mp4（H.264编码） * 分辨率：1080P、2K、4K（2K、4K人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景） * 比例：16：9或9：16 * 帧率：25 fps * 色彩标准：SDR-Rec.709（非HDR）	方案一： * 数量：1个 * 大小：3GB以内 * 时长：10秒 – 4分钟闭嘴视频 * 格式：webm（VP8、VP9编码） * 分辨率：1080P、2K、4K（2K、4K人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景） * 比例：16：9或9：16 * 帧率：25 fps * 色彩标准：SDR-Rec.709（非HDR）	方案二： * 数量：1个 * 大小：3GB以内 * 时长：10秒 – 4分钟闭嘴视频 * 格式：mp4（H.264编码） * 分辨率：1080P、2K、4K（2K、4K人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景） * 比例：16：9或9：16 * 帧率：25 fps * 色彩标准：SDR-Rec.709（非HDR）
唇动视频规格	* 数量：1个 * 大小：3GB以内 * 时长：4 – 8 分钟张嘴说话视频 * 格式：mp4（H.264编码） * 帧率：25 fps * 色彩标准：SDR-Rec.709（非HDR） * 格式、分辨率、比例：必须与底板视频的 参数保持一致 * 视频素材里面需要带有相同长度的音频	* 数量：1个 * 大小：3GB以内 * 时长：4 – 8 分钟张嘴说话视频 * 格式：webm（vp8、vp9码） * 帧率：25 fps * 色彩标准：SDR-Rec.709（非HDR） * 格式、分辨率、比例：必须与底板视频的 参数保持一致 * 视频素材里面需要带有相同长度的音频	* 数量：1个 * 大小：3GB以内 * 时长：4 – 8 分钟张嘴说话视频 * 格式：mp4（H.264编码） * 帧率：25 fps * 色彩标准：SDR-Rec.709（非HDR） * 格式、分辨率、比例：必须与底板视频的 参数保持一致 * 视频素材里面需要带有相同长度的音频
mask视频规格	无需上传	无需上传	* 数量：1个 * 大小：3GB以内 * 时长、格式、分辨率、比例、帧率：与底板视频相同

音色克隆和查询接口

🔗 接口说明

接口调用域名：<https://open.xiling.baidu.com>

接口包括提交音色克隆任务，查询克隆任务，查询克隆音色和系统公共音色列表。

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

调用本接口需要在平台购买音色克隆组件商品并绑定到应用，操作说明参考：[平台操作指引](#)

🔗 音色克隆素材要求

音频要求	1. 上传的音频文件格式需为：mp3、m4a、wav格式； 2. 上传的音频文件的时长最少应不低于30秒，最长应不超过5分钟； 3. 上传的音频文件大小需不超过20MB。
录制环境	录音环境的选择主要考虑降低噪音和混响，建议使用10平方米以内的小房间进行录音，有吸音装置的房间是理想环境。可使用价格低廉的吸音棉对房间进行改造，目的是为了改变声波的平面反射为漫反射，降低混响。 1. 尽量在同一声学环境下录制，避免过于喧哗的背景音和噪音，音频非拼接而成。 2. 避免远场环境下录制，建议离麦克风50cm之内,并需要保持位置、距离稳定，防止音量大小不一致。不要让讲话的气流正对麦克风，防止喷麦造成的噪音。 3. 尽量在自然对话状态下录制,对比刻意压低声线/呢喃获得的音色会更贴合本人。
音频质量	1. 必须仅含有一个人在说话，避免多个人同时说话。 2. 说话人发音及音质越清晰、复刻质量越高。如说话人喉音/声音底噪严重，则复刻音频喉音和底噪将会较为严重。 3. 录音时需要保持音量大小、语速稳定、注意断句、避免口腔噪音(如口水声)、杂音、混响等情况。结合最终使用场景与人设加入适当演绎，避免朗读风格过于严重，并保持整体风格一致。 4. 录音时尽量保持姿态稳定，减少不必要的肢体动作，椅子的响声、衣物的摩擦音、鼠标键盘的敲击声都是常见的人为噪声，需要格外注意细节。 5. 如口误无需终止录音，可停顿1~2秒后，继续录制即可。
录制内容	建议在录音前熟悉文案，并确定好人设及演绎风格。尽量贴近复刻内容和主题，避免说话人韵律过于平淡，否则将得到较平的复刻效果。如需要复刻情绪韵律起伏较强的音频，请尽量输入表现力强的音频。尽量不要尝试复刻小孩或者老人的音频。

🔗 接口列表

🔗 发起克隆任务

🔗 接口说明

上传音频文件，发起一个克隆任务。发起后接收任务结束的回调通知，或使用音色ID轮询任务状态（不建议）。

通过该接口克隆的音色，在中文上有更好的表现，暂不支持多语种输出。

🔗 接口协议

POST /api/digitalhuman/open/v1/tts/clone/v2

Content-Type: application/json;charset=utf-8

🔗 请求参数

参数	类型	必填	说明
name	String	是	克隆音色的名字，长度不超过 50，当前账户下需唯一
describe	String	否	克隆音色的描述，长度不超过 100
gender	String	否	性别 (male / female / unknown)
uploadAudioId	String	是	用于训练的音频文件ID，使用 文件上传接口 接口上传后获取文件ID，providerType 参数固定填写 OPEN_TTS_CLONE 1.时长在 30 秒到 5 分钟 2.大小低于 20MB 3.格式为 mp3/wav/m4a
exampleText	String	是	试听音频的文本，长度不超过 100
callbackUrl	String	否	接口调用方接受任务回调通知的url，规则和接口说明参考 接口通用说明 ，回调消息定义见下面的回调接口参数说明。

返回参数

参数	类型	必填	说明
perId	String	是	克隆音色的id

错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
14001	内部服务异常，请稍后再试
14002	内部服务异常，网络拥堵
10006	参数校验不通过
23001	语音服务内部调用异常，需联系平台解决
23003	克隆音色名字已存在

请求示例


```
{
  "requestId": "589e3dc4-0d78-48a5-8471-cf262daf8e67",
  "name": "mockname6",
  "describe": "thisisamockdesc",
  "gender": "male",
  "uploadAudioId": "617913c8-6b3c-4d30-8853-e2d0f8309768",
  "exampleText": "您好，欢迎来到曦灵数字人平台，我可以24小时在线直播",
  "callbackUrl": "http://callback/path"
}
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "perId": "10000079"
  },
  "requestId": "589e3dc4-0d78-48a5-8471-cf262daf8e67",
  "success": true
}
```

回调接口参数

规则和接口说明参考[接口通用说明](#)，回调消息定义如下：

- type: TTS_CLONE
- data:

参数名称	类型	必填	说明
perId	String	是	克隆音色的id
name	String	是	名称
gender	String	否	性别信息 (male / female / unknown)
describe	String	否	描述信息
status	String	是	状态： SUCCESS(克隆成功), FAIL(克隆失败)
exampleText	String	否	示例文本
exampleAudioUrl	String	否	示例音频URL
reason	String	否	失败原因，如果状态为失败则会出现此字段

查询音色克隆任务列表

接口说明

查询音色克隆任务列表。

接口协议

GET /api/digitalhuman/open/v1/tts/clone?isSuccess=false

请求参数

参数	类型	必填	说明
isSuccess	boolean	是	是否只查询克隆成功的任务(true: 只查询成功的任务， false: 查询全部克隆任务)
perId	String	否	查询指定 perId 的克隆任务

返回参数

参数名称	类型	必填	说明
perId	String	是	克隆音色的id
name	String	是	名称
describe	String	是	描述
exampleText	String	是	示例文本
exampleAudioUrl	String	是	示例音频URL
status	String	是	当前状态：PREPARING(准备中), CLONING(克隆中), SUCCESS(克隆成功), FAIL(克隆失败)
reason	String	是	失败原因，如果状态为 FAIL 则会出现此字段
version	String	是	音色版本，如 V2
gender	String	是	性别信息 (male / female / unknown)

请求示例

```
GET /api/digitalhuman/open/v1/tts/clone?isSuccess=false
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": [
    {
      "perId": "10000079",
      "name": "mockname",
      "describe": "thisisamockdesc",
      "exampleText": "您好, 欢迎来到曦灵数字人平台, 我可以24小时在线直播",
      "exampleAudioUrl": "",
      "status": "FAIL",
      "reason": "accountId(73894c599b274247a087fed65727e04a),
userId(73894c599b274247a087fed65727e04a) 配额余额不足",
      "version": "V2",
      "gender": "male"
    },
    {
      "perId": "10000070",
      "name": "test",
      "describe": "this is a mock desc",
      "exampleText": "您好, 欢迎来到曦灵数字人平台, 我可以24小时在线直播",
      "exampleAudioUrl": "",
      "status": "FAIL",
      "reason": "ConvertToMono failed,error:exit status 251",
      "version": "V2",
      "gender": "male"
    }
  ],
  "requestId": "d2404d3b-81fb-4cd3-81ee-4feb926fd5d1",
  "success": true
}
```

🔗 查询音色列表

🔗 接口说明

查询音色，包括克隆音色和系统库提供的公共音色。

🔗 接口协议

GET /api/digitalhuman/open/v1/tts/persons?isSystem=true

🔗 请求参数

参数	类型	必填	说明
isSystem	boolean	否	是否只查询系统音色： - true: 只查询系统音色 - false: 只查询克隆音色 - 不传：查询所有可用音色

🔗 返回参数

参数名称	类型	必填	说明
perId	String	是	克隆音色的id
name	String	是	音色名称
describe	String	否	描述
version	String	是	音色版本，如 V2
exampleAudioUrl	String	否	示例音频 URL
gender	String	否	性别信息 (male / female / unknown)
systemProvided	boolean	是	是否是系统音色

请求示例

```
GET /api/digitalhuman/open/v1/tts/persons?isSystem=true
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": [
    {
      "perId": "10001251",
      "name": "test123",
      "describe": "thisisamockdesc",
      "version": "V2",
      "exampleAudioUrl": null,
      "gender": "female",
      "systemProvided": true
    },
    {
      "perId": "100051",
      "name": "自信活泼小姐姐",
      "describe": "青年，开朗活泼",
      "version": "V2",
      "exampleAudioUrl": null,
      "gender": "female",
      "systemProvided": true
    }
  ],
  "requestId": "d2404d3b-81fb-4cd3-81ee-4feb926fd5d1",
  "success": true
}
```

2D小样本数字人形象定制-通用口型接口

接口说明

- 调用形象定制任务接口，提交真人拍摄视频，返回人像ID
- 接收任务结束的回调通知，或使用人像ID查询人像定制任务状态
- 将定制人像用于视频合成、交互等场景

接口调用域名：<https://open.xiling.baidu.com>

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

调用本接口需要在平台购买 2D 小样本形象克隆组件商品并绑定到应用，操作说明参考：[平台操作指引](#)

[2D小样本数字人（通用口型）拍摄指南](#)

🔗 接口列表

🔗 提交形象定制任务

🔗 接口说明

创建数字人训练任务，返回人像ID

🔗 接口协议

POST /api/digitalhuman/open/v1/figure/lite2d/train

Header Content-Type: application/json;charset=utf-8

🔗 请求参数

参数	类型	必填	说明
name	string	是	数字人名称 命名要求： 1. 账户内名称唯一 2. 仅可输入中文、字母、数字、中划线、下划线 3. 不超过20个字符
customizeType	string	是	定制类型，固定填：LITE_2D_GENERAL
gender	string	是	性别 1. MALE（男） 2. FEMALE（女）
keepBackground	boolean	是	是否保留底板视频背景 1. true：保留原背景，后续驱动渲染不支持替换背景图，传参说明： （1）【底板视频（templateVideoid）】上传mp4格式的视频，【mask视频（maskVideoid）】无需上传 2. false：透明背景（后续驱动渲染支持替换任意背景图），传参说明： （1）方案一（推荐）：【底板视频（templateVideoid）】上传带alpha通道webm格式的视频，【mask视频（maskVideoid）】无需上传 （2）方案二：【底板视频（templateVideoid）】和【mask视频（maskVideoid）】均需上传mp4格式的视频
templateVideoid	string	是	底板视频素材文件 ID，使用 文件上传接口 上传视频文件获取文件ID，其中 providerType 参数填写：OPEN_CUSTOMIZATION_2D_GENERAL 视频要求： 1. 数量：1个 2. 大小：5GB以内 3. 时长：10秒 - 4分钟 闭嘴视频 4. 格式：mp4（H.264编码）或 webm（vp8、vp9编码） 5. 分辨率：1080p、2k、4k（2k、4k人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景） 6. 比例：16：9 或 9：16 7. 帧率：25 fps 8. 色彩标准：SDR-Rec.709（非HDR） 2D小样本数字人（通用口型）拍摄指南
maskVideoid	string	否	底板视频对应的mask视频（人像遮罩视频）文件 ID， mask视频处理说明 ，使用 文件上传接口 接口上传视频文件获取文件ID，其中 providerType 参数填写：OPEN_CUSTOMIZATION_2D_GENERAL 若采用【透明背景类型数字人-方案二】，本参数必填。 视频要求： 1. 数量：1个 2. 大小：3GB以内 3. 格式：mp4（H.264编码） 4. 分辨率、帧率、帧数、素材内容：必须与底板视频的参数保持一致
callbackUrl	string	否	接口调用方接受任务回调通知的url，规则和接口说明参考 接口通用说明 ，回调消息定义如下表。

🔗 回调参数

- type: FIGURE_2D_LITE_GENERATE
- data:

参数名	类型	是否必填	描述
figureId	string	是	人像ID
status	string	是	状态 1. WAIT（排队中） 2. GENERATING（开始训练） 3. SUCCESS（成功） 4. FAILED（失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息

返回参数

参数	类型	必填	说明
figureId	string	是	人像ID，用于调用人像查询接口查询人像状态

请求示例

```
{
  "name": "数字人test-dong30",
  "customizeType": "LITE_2D_GENERAL",
  "gender": "MALE",
  "keepBackground": false,
  "templateVideoId": "30085e2d-b343-4b09-af25-a28430598230",
  "maskVideoId": null,
  "callbackUrl": "http://api-gateway:8080/api/digitalhuman/internal/api-gateway/v1/callback/vp/test"
}
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "figureId": "2153"
  },
  "requestId": "b978a051-fafe-4623-bfa5-1c07566e88ac",
  "success": true
}
```

错误码

错误码	描述
10006	参数校验不通过
10011	没有购买形象定制的组件商品
21001	形象包训练任务提交异常
21002	找不到人像ID
21003	人像查询异常
50001	额度预扣除失败
50002	余额不足
1002100000	情况1：透明背景，并且用的是方案一，那么有可能是webm视频没有带alpha通道
1004000000	情况1：透明背景，并且用的是方案二，那么有可能是底板视频和mask视频传反了，无法识别人脸
1006000000	情况1：透明背景，并且用的是方案二，那么有可能是底板视频和mask视频传反了，无法识别人脸
1012000000	情况1：视频开头或结尾存在黑帧（一般由剪辑软件引入），请检查删掉，并重新定制 情况2：视频逐帧读取失败，请检查视频编码是否正常，并重新定制 情况3：人脸偏转角度较大，人脸丢失，请重新拍摄 情况4：底板视频和mask视频帧数不一致
其他	内部异常请联系客服

🔗 查询人像

🔗 接口说明

分页查询定制的数字人人像列表，接口鉴权和通用分页参数查阅：[接口通用说明](#)。

🔗 接口协议

GET /api/digitalhuman/open/v1/figure/lite2d/query

🔗 GET 请求参数

参数	类型	必填	说明
figureId	string	否	指定人像ID查询，为空则查询该账号下所有人像
systemFigure	boolean	否	是否查询平台的公共人像 1. 为空：查询全部 2. true：返回平台公共人像 3. false：返回定制人像
trainSuccess	boolean	否	是否查询训练完成，状态可用的人像 1. 为空：不进行过滤 2. true：只返回可用人像（平台公共人像+训练成功状态的定制人像） 3. false：只返回排队中、训练中或训练失败的定制人像
pageNo	int	否	页码，默认为1
pageSize	int	否	每页容量，默认10

🔗 返回参数

参数	类型	必填	说明
figureId	string	是	人像ID
name	string	是	数字人名称
customizeType	string	是	定制类型
systemFigure	boolean	是	是否是平台公共人像
gender	string	是	性别
keepBackground	boolean	是	是否保留底板视频背景
taskSubmitTime	string	是	任务提交时间
templateImg	string	是	数字人缩略图
templateVideoUrl	string	是	底板视频URL
resolutionWidth	int	是	底板视频宽
resolutionHeight	int	是	底板视频高
maskVideoUrl	string	否	mask视频URL
status	string	是	状态 1. LINE_UP（排队中） 2. GENERATING（训练中） 3. SUCCESS（训练成功） 4. FAILED（训练失败）
failedCode	int	否	失败错误码
failedMessage	string	否	制作失败原因

🔗 请求示例

```
GET /api/digitalhuman/open/v1/figure/lite2d/query?figureId=931
```

🔗 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "pageNo": 1,
    "pageSize": 9,
    "totalCount": 1,
    "result": [
      {
        "figureId": "931",
        "name": "沐芸",
        "customizeType": "LITE_2D_PERSONAL",
        "systemFigure": true,
        "gender": "FEMALE",
        "keepBackground": true,
        "templateVideoUrl": "https://digital-human-pipeline-output.cdn.bcebos.com/2023-06-09T07:21:17.334022/011+1.mp4",
        "maskVideoUrl": null,
        "resolutionWidth": 1080,
        "resolutionHeight": 1920,
        "status": "SUCCESS",
        "templateImg": "http://star-light-lite.cdn.bcebos.com/2023-06-09T07:21:27.1936862023-06-09T07:21:27.193686.png",
        "taskSubmitTime": "2035-12-29 23:57:00",
        "failedCode": 0,
        "failedMessage": null
      }
    ]
  },
  "requestId": "ac0c9acd-69fa-40bd-aa37-f2463beb2f30",
  "success": true
}
```

2D小样本数字人形象定制-专属口型接口

🔗 接口说明

接口调用域名：<https://open.xiling.baidu.com>

- 1. 调用形象定制任务接口，提交真人拍摄视频，返回人像ID
- 2. 接收任务结束的回调通知，或使用人像ID查询人像定制任务状态
- 3. 将定制人像用于视频合成、交互等场景

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

调用本接口需要在平台购买 2D 小样本形象克隆组件商品并绑定到应用，操作说明参考：[平台操作指引](#)

[2D小样本数字人（专属口型）拍摄指南](#)

🔗 接口列表

🔗 提交形象定制任务

🔗 接口说明

创建数字人训练任务，返回人像ID

🔗 接口协议

POST /api/digitalhuman/open/v1/figure/lite2d/train

Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必填	说明
name	string	是	数字人名称 命名要求： 1. 账户内名称唯一 2. 仅可输入中文、字母、数字、中划线、下划线 3. 不超过20个字符
customizeType	string	是	定制类型，固定填：LITE_2D_PERSONAL
gender	string	是	性别 1. MALE（男） 2. FEMALE（女）
keepBackground	boolean	是	是否保留底板视频背景 1. true：固定背景（保留原背景，后续驱动渲染不支持替换背景图） （1）【底板视频（templateVideoId）】【唇动视频（lipVideoId）】上传mp4格式的视频，【底板mask视频（maskVideoId）】无需上传 2. false：透明背景（后续驱动渲染支持替换任意背景图） （1）方案一（推荐）：【底板视频（templateVideoId）】【唇动视频（lipVideoId）】上传带alpha通道webm格式的视频，【底板mask视频（maskVideoId）】无需上传 （2）方案二：【底板视频（templateVideoId）】【唇动视频（lipVideoId）】【底板mask视频（maskVideoId）】均需上传mp4格式的视频
templateVideoId	string	是	底板视频素材文件 ID，使用 文件上传接口 上传视频文件获取文件ID，其中 providerType 参数填写：OPEN_CUSTOMIZATION_2D_EXCLUSIVE 视频要求： 1. 数量：1个 2. 大小：5GB以内 3. 时长：10秒 - 8分钟 闭嘴视频 4. 格式：mp4（H.264编码）或 webm（vp8、vp9码） 5. 分辨率：1080p、2k、4k（2k、4k人像训练限时免费，只支持用于视频合成场景，不支持交互、直播等实时渲染推流场景） 6. 比例：16：9 或 9：16 7. 帧率：25 fps 8. 色彩标准：SDR-Rec.709（非HDR） 9. 视频素材里面需要带有相同长度的音频 2D小样本数字人（专属口型）拍摄指南
lipVideoId	string	是	唇动视频素材文件 ID，使用 文件上传接口 上传视频文件获取文件ID，其中 providerType 参数填写：OPEN_CUSTOMIZATION_2D_EXCLUSIVE 视频要求： 1. 数量：1个 2. 大小：5GB以内 3. 时长：4 - 8 分钟 张嘴说话视频 4. 格式：mp4（H.264编码）或 webm（vp8、vp9码） 5. 帧率：25 fps 6. 色彩标准：SDR-Rec.709（非HDR） 7. 格式、分辨率、比例： 必须与底板视频的参数保持一致 8. 视频素材里面需要带有相同长度的音频

			8. 视频素材里面需要带有相同长度的音频 2D小样本数字人（专属口型）拍摄指南
maskVideoId	string	否	底板视频对应的mask视频（人像遮罩视频）文件 ID， mask视频处理说明 ，使用 文件上传接口 接口上传视频文件获取文件ID，其中 providerType 参数填写： OPEN_CUSTOMIZATION_2D_EXCLUSIVE 若采用【透明背景类型数字人-方案二】，本参数必填。 视频要求： 1. 数量：1个 2. 大小：3GB以内 3. 格式：mp4（H.264编码） 4. 素材内容、分辨率、帧率、帧数：必须与底板视频的参数保持一致
callbackUrl	string	否	接口调用方接受任务回调通知的url，规则和接口说明参考 接口通用说明 ，回调消息定义如下表。

回调参数

- type: FIGURE_2D_LITE_GENERATE
- data:

参数名	类型	是否必填	描述
figureId	string	是	人像ID
status	string	是	状态 1. WAIT（排队中） 2. GENERATING（开始训练） 3. SUCCESS（成功） 4. FAILED（失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息

返回参数

参数	类型	必填	说明
figureId	string	是	人像ID，用于调用人像查询接口查询人像状态

请求示例

```
{
  "name": "数字人test-dong30",
  "customizeType": "LITE_2D_GENERAL",
  "gender": "MALE",
  "keepBackground": false,
  "templateVideoId": "30085e2d-b343-4b09-af25-a28430598230",
  "maskVideoId": null,
  "callbackUrl": "http://api-gateway:8080/api/digitalhuman/internal/api-gateway/v1/callback/vp/test"
}
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "figureId": "2153"
  },
  "requestId": "b978a051-fafe-4623-bfa5-1c07566e88ac",
  "success": true
}
```

错误码

错误码	描述
10006	参数校验不通过
10011	没有购买形象定制的组件商品
21001	形象包训练任务提交异常
21002	找不到人像ID
21003	人像查询异常
50001	额度预扣除失败
50002	余额不足
1002100000	情况1：透明背景，并且用的是方案一，那么有可能是webm视频没有带alpha通道
1004000000	情况1：透明背景，并且用的是方案二，那么有可能是底板视频和mask视频传反了，无法识别人脸
1006000000	情况1：透明背景，并且用的是方案二，那么有可能是底板视频和mask视频传反了，无法识别人脸
1012000000	情况1：视频开头或结尾存在黑帧（一般由剪辑软件引入），请检查删掉，并重新定制 情况2：视频逐帧读取失败，请检查视频编码是否正常，并重新定制 情况3：人脸偏转角度较大，人脸丢失，请重新拍摄 情况4：底板视频和mask视频帧数不一致
其他	内部异常请联系客服

查询人像

接口说明

分页查询定制的数字人人像列表，接口鉴权和通用分页参数查阅：[接口通用说明](#)。

接口协议

GET /api/digitalhuman/open/v1/figure/lite2d/query

GET 请求参数

参数	类型	必填	说明
figureId	string	否	指定人像ID查询，为空则查询该账号下所有人像
systemFigure	boolean	否	是否查询平台的公共人像 1. 为空：查询全部 2. true：返回平台公共人像 3. false：返回定制人像
trainSuccess	boolean	否	是否查询训练完成，状态可用的人像 1. 为空：不进行过滤 2. true：只返回可用人像（平台公共人像+训练成功状态的定制人像） 3. false：只返回排队中、训练中或训练失败的定制人像
pageNo	int	否	页码，默认为1
pageSize	int	否	每页容量，默认10

返回参数

参数	类型	必填	说明
figureId	string	是	人像ID
name	string	是	数字人名称
customizeType	string	是	定制类型
systemFigure	boolean	是	是否是平台公共人像
gender	string	是	性别
keepBackground	boolean	是	是否保留底板视频背景
taskSubmitTime	string	是	任务提交时间
templateImg	string	是	数字人缩略图
templateVideoUrl	string	是	底板视频URL
resolutionWidth	int	是	底板视频宽
resolutionHeight	int	是	底板视频高
maskVideoUrl	string	否	mask视频URL
status	string	是	状态 1. LINE_UP（排队中） 2. GENERATING（训练中） 3. SUCCESS（训练成功） 4. FAILED（训练失败）
failedCode	int	否	失败错误码
failedMessage	string	否	制作失败原因

请求示例

GET /api/digitalhuman/open/v1/figure/lite2d/query?figureId=931

返回示例


```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "pageNo": 1,
    "pageSize": 9,
    "totalCount": 1,
    "result": [
      {
        "figureId": "931",
        "name": "沐芸",
        "customizeType": "LITE_2D_PERSONAL",
        "systemFigure": true,
        "gender": "FEMALE",
        "keepBackground": true,
        "templateVideoUrl": "https://digital-human-pipeline-output.cdn.bcebos.com/2023-06-09T07:21:17.334022/011+1.mp4",
        "maskVideoUrl": null,
        "resolutionWidth": 1080,
        "resolutionHeight": 1920,
        "status": "SUCCESS",
        "templateImg": "http://star-light-lite.cdn.bcebos.com/2023-06-09T07:21:27.1936862023-06-09T07:21:27.193686.png",
        "taskSubmitTime": "2035-12-29 23:57:00",
        "failedCode": 0,
        "failedMessage": null
      }
    ]
  },
  "requestId": "ac0c9acd-69fa-40bd-aa37-f2463beb2f30",
  "success": true
}
```

端渲染2D小样本数字人拍摄指南

定制端渲染2D小样本数字人时，我们需要您提供一段10秒-4分钟的闭嘴动作底板视频，和一段3-4分钟的真人口播唇动视频，如需定制人像，请通过[工单](#)联系我们。

为了提升最终生成的数字人效果，我们特别准备了以下拍摄指南，供您参考：

示例视频部分片段：

绿幕拍摄：训练透明背景数字人（可替换任意背景）

- 底板视频：一段10秒-4分钟的闭嘴视频，用于人像动作训练，训练出的数字人动作会与该底板完全一致

<https://meta-human-editor-prd.cdn.bcebos.com/91eacc90-a5cc-444b-997c-d33afdcada8f/ad42f40b-96ce-4e78-ac2f-b0264053108e/底板视频.mp4>

- 唇动视频：一段3-4分钟的真人口播视频，用于数字人口型训练，视频需保留音轨，音画同步，模特声音清晰无环境噪音

<https://meta-human-editor-prd.cdn.bcebos.com/91eacc90-a5cc-444b-997c-d33afdcada8f/4be7261c-04e2-4459-88dd-ecfac829bccd/唇动视频2.mp4>

🔗 1 拍摄准备

1. 模特服饰

- 可按照模特喜好选择，与背景颜色风格协调、适合实际应用场景即可

- 合身，不能过大或过紧
- 整洁大方，无明显褶皱
- 穿低领衣服，露出脖子
- 避免碎花、格纹、条纹等大面积花纹的衣服，避免摩尔纹产生
- 建议着深色服装，或者选择其他与背景颜色反差大的服饰颜色
- 不建议选择白色、反光材质、过于光滑面料，会反绿（绿幕拍摄需注意）
- 服饰中不存在绿色系服饰，不存在纱质材料，会漏绿（绿幕拍摄需注意）

2. 模特妆造

- 妆容自然大方，有亲和力，无明显突兀感
- 面部白皙且有立体感，但不宜有过重阴影
- 如果存在瑕疵，例如皮肤不平整，长痘，手部存在血丝，需要遮瑕
- 嘴唇过薄可通过口红补救
- 露出双耳，可佩戴固定耳饰，不建议佩戴易晃动耳饰
- 注意手部、脖子、腿等部位，需要保持与面部肤色一致
- 头发正面看无碎发、无毛刺、无镂空，饱满有弧度（绿幕拍摄需注意）

3. 播报文本：

- 准备3-4分钟约800字以上可读性较强的语料
- 文本内容最好贴合实际应用场景
- 建议模特在录制前熟读文本

4. 布景：

- 实景拍摄：背景整洁、无密集纹理，适合实际应用场景
- 绿幕拍摄：绿幕颜色统一，且平整无褶皱；建议人物和绿幕距离大于2米，避免人物反绿

5. 灯光：

- 用灯光从斜前侧方照亮模特面部，使主播在画面中更突出
- 在灯光下，模特皮肤白皙透亮，避免脸部反光、肤色暗沉或曝光
- 模特光线均匀，尤其注意脸部，颈部，手部，腿部的光照一致性
- 背景光照均匀，并与人像轮廓形成清晰边界
- 灯光一旦确定，切勿再次调整，否则全部重拍

6. 设备：

- 建议拍摄的分辨率不低于1080P，关闭HDR模式，关闭log模式
- 需准备提词器，放映预先准备的播报文本
- 比例：竖屏为9:16
- 角度：镜头高度与主播眼睛保持同一水平线，模特在头部保持自然的情况下平视镜头，建议正对镜头，侧身不超过10度，不要仰头、低头、仰视、俯视

- 构图：将主播放置于画面中间，确保五官清晰、面部无遮挡
- 可使用三脚架或稳定器固定拍摄设备，以减少画面抖动，确保视频流畅稳定

🔗 2 形象拍摄

🔗 底板视频：

1. 面部神情：头部保持相对稳定，模仿说话状态轻微晃动，面带微笑，面部肌肉放松，注意眼睛一直看镜头，避免张口露齿、抿嘴、眼睛四处看（无需说话）
2. 模特动作：越自然越好，以下建议仅供参考
 - 模特面带微笑，头部微微自然晃动、身体稳定、手部自然放在身前，偶尔左右手交替/手放下再端起即可，不需要做过多动作，因为交互场景某些时候数字人处于聆听静默状态，这时数字人不需要做过多动作，生成的数字人姿态动作将与视频完全一致，请确保拍摄时的姿态可用于后续内容创作
 - 手势动作避免遮挡嘴巴、面部、脖子周围，也避免出画面

🔗 唇动视频：

1. 唇动视频模特形象、机位、灯光等需与底板视频完全一致，建议模特连续拍摄底板视频与唇动视频，后期剪辑
2. 在句与句之间多停顿，停顿时自然闭嘴

tips：

- 可以连续拍摄10分钟左右，为剪辑提供更多的素材

🔗 3 后期处理

🔗 3.1 剪辑-调色美颜

- 需**掐头去尾剪辑**，最终提供的底板视频需全程无拼接，跳帧对视频效果影响较大
- 开头保留2-3秒没有手势的静默状态，可以使动作自然开始
- 片段删减说明
 - 不建议保留：不通用动作、指向性动作、否定动作、眼神乱瞟、笑场等不合时宜的动作，有较大噪音、其他人说话等片段
 - 可以保留：录制过程中微小卡顿，偶尔念错对训练效果基本没影响

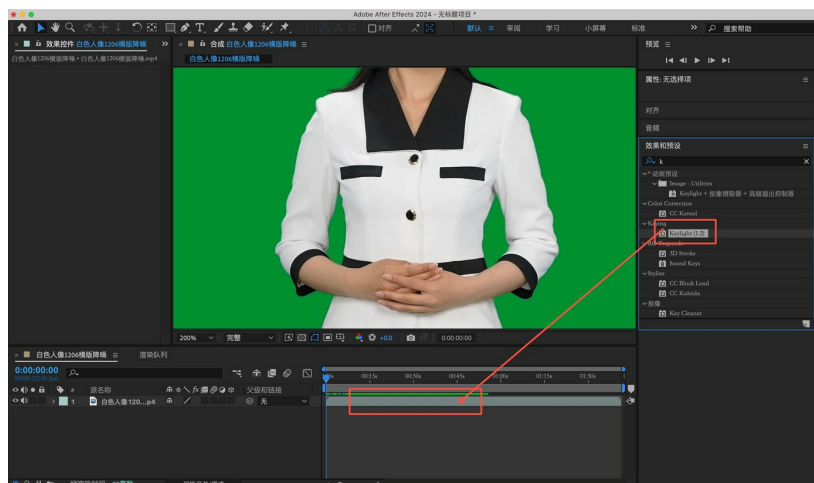
如果主播有美颜需求，也可以在剪辑时用美颜软件还原美貌，数字人训练效果会与视频素材保持一致

- 若无需抠绿，可直接导出视频，实景拍摄的底板视频规格要求如本文【4.2 视频规格要求】

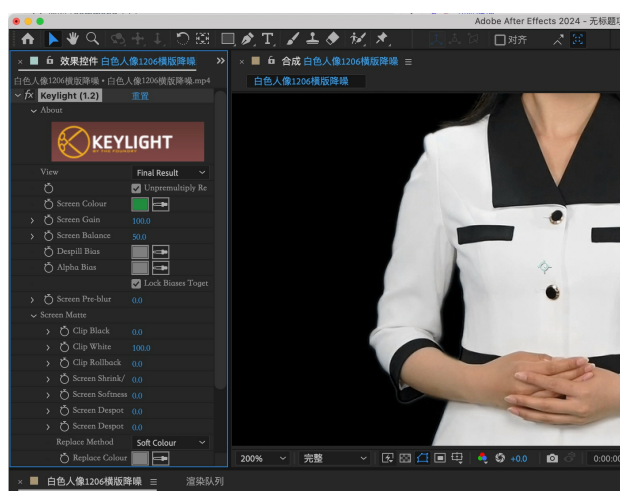
🔗 3.2 抠绿（绿幕拍摄）

以下为AE（Adobe After Effects）抠绿方式及导出视频规格，也可选择其他软件进行抠绿处理。

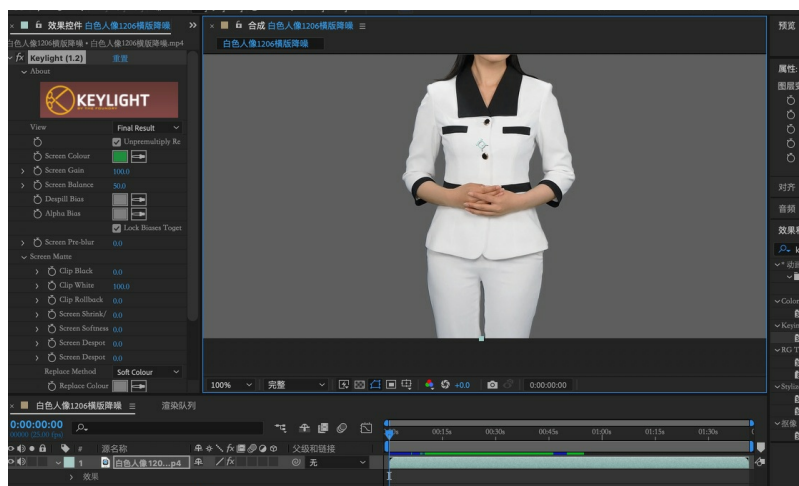
1. 对视频添加超级键效果：



2. 调整各种参数：抠除绿色背景，去除模特脸部以及身上绿色反光

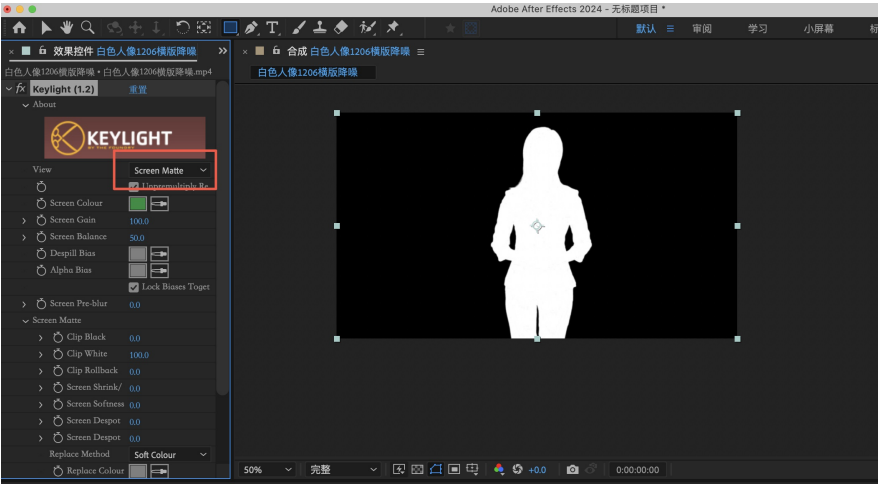


3. 检查抠绿效果：抠绿后可换浅灰色背景，检查抠绿是否干净以及是否有该保留的部分被抠除



4. 导出视频：

导出mask视频，选择【mask模式】方式如下，视频规格要求如本文【4.2 视频规格要求】



4 视频自检

在提交视频前，请务必逐条确认视频素材是否满足以下要求。

4.1 内容要求：

- 动作自然，无指向性动作
- 在句与句之间多停顿，停顿时自然闭嘴
- 头部无大角度转动，脸部全程无遮挡
- 提供的视频全程无拼接、无跳帧
- 视频音画同步，无明显噪音
- 人像边缘平滑自然，没有绿色残留或锯齿状边缘（若抠绿需注意）
- 穿低领衣服，露出脖子
- 光线充足，脸部整体光线均匀
- 保持妆造、光线、景别不变

4.2 视频规格要求：

	固定背景（保留原背景）	透明背景（可替换任意背景）
拍摄布景	实景拍摄	绿幕拍摄
底板视频规格	* 数量：1个 * 大小：700M以内 * 时长：10秒 - 4分钟闭嘴视频 * 格式：mp4（H.264编码） * 分辨率：1080P * 比例：竖屏9：16 * 帧率：25 fps	* 数量：1个 * 大小：700M以内 * 时长：10秒 - 4分钟闭嘴视频 * 格式：mp4（H.264编码） * 分辨率：1080P * 比例：竖屏9：16 * 帧率：25 fps
层动视频规格	* 数量：1个 * 大小：700M以内 * 时长：3-4分钟张嘴说话视频 * 格式：mp4（H.264编码） * 帧率：25 fps * 格式、分辨率、比例：必须与底板视频的参数保持一致 * 视频素材里面需要带有相同长度的音频	* 数量：1个 * 大小：700M以内 * 时长：3-4分钟张嘴说话视频 * 格式：mp4（H.264编码） * 帧率：25 fps * 格式、分辨率、比例：必须与底板视频的参数保持一致 * 视频素材里面需要带有相同长度的音频
mask视频规格	无需上传	* 数量：1个mask视频 * 大小：700M以内 * 时长、格式、分辨率、比例、帧率：与底板视频相同

5 大屏人像素材比例要求

确定用于大屏的人像可参考本素材规范，保证人像在各尺寸数字人大屏上展示效果较佳，防止人像上大屏出现比例不一致、遮挡卡片等现象。

	43寸半身屏（1080*1920分辨率）	72、75寸全身屏（分别宽1770、2160px，高3840px）
人像素材要求	- 人像头顶至素材底部，参考最佳高度：1800px（占比93.8%）参考最小高度：1750px（占比91.1%） - 人像下巴至素材底部，参考最佳高度：1390px（占比72.4%）参考最小高度：1340px（占比69.8%） - 人像肩宽，最佳肩宽650px（占比60.2%）	- 人像头顶至素材底部，参考高度：3580px（占比93.2%）。 - 人像下巴至素材底部，参考高度：3130px（占比81.5%）。 - 人像脚底至素材底部，参考高度：100px（占比2.6%）。
图示		

2D小样本数字人抠像接口

接口说明

接口调用地址：<https://open.xiling.baidu.com>

抠像接口支持对用户上传的视频进行背景扣除处理。该接口能够智能识别并保留视频中的人像，无论是实景视频还是绿幕拍摄，都可以实现高效、精准的背景去除。适用于需要背景替换或去除的各种应用场景。

本功能需要应用绑定「形象定制-2D小样本数字人（通用口型）」或者「形象定制-2D小样本数字人（专属口型）」组件才能使用。

拍摄指南参考：https://cloud.baidu.com/doc/AI_DH/s/8m0uiiv0m

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

接口列表

抠像任务

接口说明

提供视频背景抠像功能，提交任务，返回抠像处理后的视频URL及视频ID，视频ID可直接用于2D小样本形象定制。

接口协议

POST /api/digitalhuman/open/v1/figure/lite2d/background/submit

Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必填	说明
videoid	string	是	视频素材文件 ID，视频素材不允许包含黑帧，否则人像训练会导致帧数不匹配报错，使用 文件上传接口 接口上传视频文件获取文件ID，其中providerType 参数填写： 通用口型人像：OPEN_CUSTOMIZATION_2D_GENERAL 专属口型人像：OPEN_CUSTOMIZATION_2D_EXCLUSIVE 视频要求： 1. 数量：1个 2. 大小：3GB以内 3. 时长：10秒-8分钟以内 4. 格式：mp4（H.264编码） 5. 分辨率：1080p、2k、4k 6. 比例：16：9或9：16 7. 帧率：25 fps 8. 色彩标准：SDR-Rec.709（非HDR）
type	string	是	视频类型： 1. GREEN：表示绿幕背景，抠图时去掉绿幕，保留前景； 2. REAL：表示实景视频，抠图只能保留人像。
callbackUrl	string	否	接口调用方接受任务回调通知的url，规则和接口说明参考 接口通用说明 ，回调消息定义如下表。

回调参数

- type: FIGURE_2D_LITE_BACKGROUND
- data:

参数名	类型	是否必填	描述
taskId	string	是	任务ID
status	string	是	状态： 1. LINE_UP（排队中） 2. GENERATING（抠像中） 3. SUCCESS（成功） 4. FAILED（失败）
videoMaskId	string	否	抠像后的视频MaskID，在人像训练接口使用（省去视频上传的步骤），当status=SUCCESS时会有值。
videoUrl	string	否	抠像后的视频url（仅用来查看效果），当status=SUCCESS时会有值。
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息

返回参数

参数	类型	必填	说明
taskId	string	是	任务ID，用于调用人像查询接口查询抠像任务状态

请求示例

```
{
  "videoid": "cddbadc3-47ad-4bc3-bb0e-be58c8e4c681",
  "type": "REAL",
  "callbackUrl": "http://api-gateway:8080/api/digitalhuman/internal/api-gateway/v1/callback/vp/test"
}
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "bkg-qmur7u3her5cn255"
  },
  "requestId": "d1f10c73-f43d-4933-ad9f-4cd20df5da55",
  "success": true
}
```

错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
21007	【视频素材】该文件不是您的，请检查
30009	获取【视频素材】异常：查询的fileId不存在
150011	抠图任务提交失败，请稍后再试
150012	抠图任务训练失败，请稍后再试
其他错误码	内部服务异常，请稍后再试或联系客服处理

查询任务

接口说明

使用任务ID查询任务状态

接口协议

GET /api/digitalhuman/open/v1/figure/lite2d/background/task

GET 请求参数

参数	类型	必填	说明
taskId	string	否	为空则查询该账号下所有任务ID
pageNo	int	否	页码，默认为1
pageSize	int	否	每页容量，默认10

返回参数

参数	类型	必填	说明
taskId	string	是	任务ID
status	string	是	状态： 1. LINE_UP（排队中） 2. GENERATING（抠像中） 3. SUCCESS（成功） 4. FAILED（失败）
videoMaskId	string	否	抠像后的视频MaskID，在人像训练接口使用（省去视频上传的步骤），当status=SUCCESS时会有值。
videoUrl	string	否	抠像后的视频url（仅用来查看效果），当status=SUCCESS时会有值。
type	string	是	视频类型 1. GREEN，表示绿幕背景，抠图时去掉绿幕，保留前景； 2. REAL，表示实景视频，抠图只能保留人像。
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
createTime	string	是	创建时间，示例：2024-04-29T20:28:06
updateTime	string	是	更新时间，示例：2024-04-29T20:28:06

请求示例

```
GET /api/digitalhuman/open/v1/figure/lite2d/background/task?taskId=bkg-qmur7u3her5cn255
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "pageNo": 1,
    "pageSize": 10,
    "totalCount": 16,
    "result": [
      {
        "taskId": "bkg-qm0vj6dmkdah047w",
        "status": "SUCCESS",
        "videoMaskId": "21ca0c1c-338a-472c-8ce4-79c84ceb4f30",
        "videoUrl": "http://bridge.bj.bcebos.com/2d_train_server_open_sanbox/bkg-qm0vj6dmkdah047w_/work/workdir/bkg-qm0vj6dmkdah047w/results/3ec90927-9ab4-4a9a-a993-7e8f382dcdde_fusion.mp4?authorization=bce-auth-v1%2Fa0208a0779694c2fbefe77194d71db41%2F2024-12-25T11%3A42%3A46Z%2F-1%2F%2F691e70579efa0d5996108729dbf855d96279e1af9fa9ec5dbc02d22e656b2bc0",
        "type": "REAL",
        "failedCode": 0,
        "createTime": "2024-12-25 19:40:20",
        "updateTime": "2024-12-25 19:42:46"
      }
    ]
  },
  "requestId": "fc7b4d7f-6d38-404b-a1c9-ff89ffe5c1f",
  "success": true
}
```

声音克隆LITE版

🔗 接口说明

LITE版新增支持多语种克隆功能，上传录制的中文音频，克隆后的音色支持输出多国语言。

接口调用域名：<https://open.xiling.baidu.com>

接口包括提交音色克隆任务，查询克隆任务，查询克隆音色和系统公共音色列表。

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

调用本接口需要在平台购买音色克隆组件商品并绑定到应用，操作说明参考：[平台操作指引](#)

音色克隆	音色克隆LITE
200元/个音色	10元/个音色
功能：上传至少30s的音频克隆音色	功能：上传至少10s的音频克隆音色
特点：音色高度还原，更懂中文表达。	特点：音色基本还原，更高性价比。

🔗 新增：支持输出多语种

支持输出语种类型	Chinese（中文）；Chinese,Yue（粤语）；English（英语）；Russian（俄语）；Spanish（西班牙语）；French（法语）；Portuguese（葡萄牙语）；German（德语）；Turkish（土耳其语）；Dutch（荷兰语）；Ukrainian（乌克兰语）；Vietnamese（越南语）；Indonesian（印尼语）；Japanese（日语）；Italian（意大利语）；Korean（韩语）；Thai（泰语）；Polish（波兰语）；Romanian（罗马尼亚语）；Greek（希腊语）；Czech（捷克语）；Finnish（芬兰语）；Hindi（印地语）
----------	--

音色克隆素材要求

• 建议参考样例音频效果 [试听](#)

音频要求	1. 上传的音频文件格式需为：mp3、m4a、wav格式； 2. 上传的音频文件的时长最少应不低于10秒，最长应不超过5分钟； 3. 上传的音频文件大小需不超过20MB
录制环境	录音环境的选择主要考虑降低噪音和混响，建议使用10平方米以内的小房间进行录音，有吸音装置的房间是理想环境。可使用价格低廉的吸音棉对房间进行改造，目的是为了改变声波的平面反射为漫反射，降低混响。1. 尽量在同一声学环境下录制，避免过于喧哗的背景音和噪音，音频非拼接而成。2. 避免远场环境下录制，建议离麦克风50cm之内,并需要保持位置、距离稳定，防止音量大小不一致。不要让讲话的气流正对麦克风，防止喷麦造成的噪音。3. 尽量在自然对话状态下录制,对比刻意压低声线/呢喃获得的音色会更贴合本人。
音频质量	1. 必须仅含有一个人在说话，避免多个人同时说话。 2. 说话人发音及音质越清晰、复刻质量越高。如说话人喉音/声音底噪严重，则复刻音频喉音和底噪将会较为严重。 3. 录音时需要保持音量大小、语速稳定、注意断句、避免口腔噪音(如口水声)、杂音、混响、回声等情况。结合最终使用场景与人设加入适当演绎，避免朗读风格过于严重，并保持整体风格一致。 4. 录音时尽量保持姿态稳定，减少不必要的肢体动作，椅子的响声、衣物的摩擦音、鼠标键盘的敲击声都是常见的人为噪声，需要格外注意细节。 5. 如口误无需终止录音，可停顿1~2秒后，继续录制即可。
录制内容	建议在录音前熟悉文案，并确定好人设及演绎风格。尽量贴近复刻内容和主题，避免说话人韵律过于平淡，否则将得到较平的复刻效果。如需要复刻情绪韵律起伏较强的音频，请尽量输入表现力强的音频。尽量不要尝试复刻小孩或者老人的音频。

接口列表

发起克隆任务

接口说明 上传音频文件，发起一个轻量级音色克隆任务。发起后接收任务结束的回调通知，或使用音色ID轮询任务状态（不建议）。

接口协议

```
POST /api/digitalhuman/open/v1/tts/clone/lite
Content-Type: application/json;charset=utf-8
```

请求参数 | 参数 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | name | String | 是 | 克隆音色的名字，长度不超过 50，当前账户下需唯一 | | describe | String | 否 | 克隆音色的描述，长度不超过 100 | | gender | String | 否 | 性别 (male / female / unknown) | | uploadAudioId | String | 是 | 用于训练的音频文件ID，使用[文件上传接口](#)接口上传后获取文件ID，providerType 参数固定填写 OPEN_TTS_CLONE_LITE | | exampleText | String | 是 | 试听音频的文本，长度不超过 100 | | callbackUrl | String | 否 | 接口调用方接受任务回调通知的url，规则和接口说明参考[接口通用说明](#)，回调消息定义见下面的回调接口参数说明。 |

返回参数 | 参数 | 类型 | 必填 | 说明 || --- | --- | --- | --- || perId | String | 是 | 克隆音色的id |

错误码 | 错误码 | 描述 || --- | --- || 0 | 正常返回 || 4911 | 找不到app信息，请确认appId是否输入正确 || 4913 | 无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用 || 10001 | 签名校验失败 || 10002 | 签名信息为空 || 10003 | 签名格式错误 || 10004 | 未识别错误的通用错误码 || 10005 | 请求体JSON解析失败，请确认是否是合法的JSON格式 || 10006 | 参数校验不通过 || 10011 | 没有购买对应的商品 || 14001 | 内部服务异常，请稍后再试 || 14002 | 内部服务异常，网络拥堵 || 10006 | 参数校验不通过 || 23001 | 语音服务内部调用异常，需联系平台解决 || 23003 | 克隆音色名字已存在 |

请求示例

```
{
  "name": "mockname6",
  "describe": "thisisamockdesc",
  "gender": "male",
  "uploadAudiold": "617913c8-6b3c-4d30-8853-e2d0f8309768",
  "exampleText": "您好，欢迎来到曦灵数字人平台，我可以24小时在线直播",
  "callbackUrl": "http://callback/path"
}
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "perId": "10000079"
  },
  "requestId": "589e3dc4-0d78-48a5-8471-cf262daf8e67",
  "success": true
}
```

回调接口参数 规则和接口说明参考[接口通用说明](#)，回调消息定义如下：

- type: TTS_CLONE_LITE
- data:

参数名称	类型	必填	说明
perId	String	是	克隆音色的id
name	String	是	名称
gender	String	否	性别信息 (male / female / unknown)
describe	String	否	描述信息
status	String	是	状态： SUCCESS(克隆成功), FAIL(克隆失败)
exampleText	String	否	示例文本
exampleAudioUrl	String	否	示例音频URL
reason	String	否	失败原因，如果状态为失败则会出现此字段

🔗 查询音色克隆任务列表

接口说明 查询轻量级音色克隆任务列表。

接口协议

```
GET /api/digitalhuman/open/v1/tts/clone?isSuccess=false
```

请求参数 | 参数 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | isSuccess | boolean | 是 | 是否只查询克隆成功的任务(true: 只查询成功的任务， false: 查询全部克隆任务) | | perId | String | 否 | 查询指定 perId 的克隆任务 |

返回参数 | 参数名称 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | perId | String | 是 | 克隆音色的id | | name | String | 是 | 名称 | | describe | String | 是 | 描述 | | exampleText | String | 是 | 示例文本 | | exampleAudioUrl | String | 是 | 示例音频URL | | status | String | 是 | 当前状态：PREPARING(准备中), CLONING(克隆中), SUCCESS(克隆成功), FAIL(克隆失败) | | reason | String | 是 | 失败原因，如果状态为 FAIL 则会出现此字段 | | version | String | 是 | 音色版本，LITE | | gender | String | 是 | 性别信息 (male / female / unknown) |

请求示例

```
GET /api/digitalhuman/open/v1/tts/clone?isSuccess=false
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": [
    {
      "perId": "10000079",
      "name": "mockname",
      "describe": "thisisamockdesc",
      "exampleText": "您好，欢迎来到曦灵数字人平台，我可以24小时在线直播",
      "exampleAudioUrl": "",
      "status": "FAIL",
      "reason": "accountId(73894c599b274247a087fed65727e04a),
      userId(73894c599b274247a087fed65727e04a) 配额余额不足",
      "version": "LITE",
      "gender": "male"
    },
    {
      "perId": "10000070",
      "name": "test",
      "describe": "this is a mock desc",
      "exampleText": "您好，欢迎来到曦灵数字人平台，我可以24小时在线直播",
      "exampleAudioUrl": "",
      "status": "FAIL",
      "reason": "ConvertToMono failed,error:exit status 251",
      "version": "V2",
      "gender": "male"
    }
  ],
  "requestId": "d2404d3b-81fb-4cd3-81ee-4feb926fd5d1",
  "success": true
}
```

🔗 查询音色列表

接口说明 查询音色，包括克隆音色和系统库提供的公共音色。

接口协议

```
GET /api/digitalhuman/open/v1/tts/persons?isSystem=true
```

请求参数 | 参数名称 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | isSystem | boolean | 否 | 是否只查询系统音色：
true:只查询系统音色
false:只查询克隆音色
不传:查询所有可用音色 |

返回参数 | 参数名称 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | perId | String | 是 | 克隆音色的id | | name | String | 是 | 音色名称 | | describe | String | 否 | 描述 | | version | String | 是 | 音色版本，如 V2 | | exampleAudioUrl | String | 否 | 示例音频 URL | | gender | String | 否 | 性别信息 (male / female / unknown) | | systemProvided | boolean | 是 | 是否是系统音色 |

请求示例

```
GET /api/digitalhuman/open/v1/tts/persons?isSystem=true
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": [
    {
      "perId": "10001251",
      "name": "test123",
      "describe": "thisisamockdesc",
      "version": "V2",
      "exampleAudioUrl": null,
      "gender": "female",
      "systemProvided": true
    },
    {
      "perId": "100051",
      "name": "自信活泼小姐姐",
      "describe": "青年，开朗活泼",
      "version": "V2",
      "exampleAudioUrl": null,
      "gender": "female",
      "systemProvided": true
    }
  ],
  "requestId": "d2404d3b-81fb-4cd3-81ee-4feb926fd5d1",
  "success": true
}
```

生成式数字人形象定制接口

🔗 接口说明

- 1. 调用生成式形象定制任务接口，提交人像图片，返回人像ID

- 2. 接收任务结束的回调通知，或使用人像ID查询人像定制任务状态
- 3. 定制的生成式人像适用于视频合成场景
- 4. 生成式数字人分为【稳定版】和【灵动版】。生成式具有不确定性，可能存在眼神外瞟、镜头移动、手部变形等情况。

【稳定版】身体呈现自然动态，手部几乎无大幅度动作，效果更稳定。[参考视频](#)

【灵动版】身体呈现自然动态，手部大概率出现摊手动作，效果更灵动。[参考视频](#)

接口调用域名：<https://open.xiling.baidu.com>

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

调用本接口需要在平台购买生成式数字人克隆组件商品并绑定到应用。

🔗 接口列表

🔗 提交形象定制任务

🔗 接口说明

创建数字人训练任务，返回人像ID

🔗 接口协议

POST /api/digitalhuman/open/v1/figure/generative/train

Header Content-Type: application/json;charset=utf-8

🔗 请求参数

参数	类型	必填	说明
name	string	是	数字人名称 命名要求： 1. 账户内名称唯一 2. 不超过20个字符
customizeType	string	是	定制类型 1. 稳定版：GENERATE_NO_ACTION 2. 灵动版：GENERATE_ACTION
gender	string	是	性别 1. MALE（男） 2. FEMALE（女）
figureImageUrl	string	是	参考图像 1. 支持传入图片Base64编码或图片URL 2. 图片格式支持.jpg / .jpeg / .png 3. 图片文件大小不能超过10MB 4. 图片分辨率不小于720*720px 5. 图片宽高比要在1:2 ~ 2:1之间
callbackUrl	string	否	接口调用方接受任务回调通知的url，规则和接口说明参考 接口通用说明 ，回调消息定义如下表。

🔗 回调参数

- type: FIGURE_2D_LITE_GENERATE
- data:

参数名	类型	是否必填	描述
figureId	string	是	人像ID
status	string	是	状态 1. LINE_UP（排队中） 2. ACTIONING（动作生成中） 3. GENERATING（开始训练） 4. SUCCESS（成功） 5. FAILED（失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息

返回参数

参数	类型	必填	说明
figureId	string	是	人像ID，用于调用人像查询接口查询人像状态

请求示例

```
{
  "name": "数字人test2025",
  "customizeType": "GENERATE_NO_ACTION", // 无动作
  "figureImageUrl": "https://open-api-test.bj.bcebos.com/%E7%94%9F%E6%88%90%E5%BC%8F%E6%B5%8B%E8%AF%95%2FSlide%2016_9%20-%201.png",
  "gender": "MALE"
}
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "figureId": "2153"
  },
  "requestId": "b978a051-fafe-4623-bfa5-1c07566e88ac",
  "success": true
}
```

错误码

错误码	描述
10006	参数校验不通过
10011	没有购买形象定制的组件商品
10020	图片内容有安全风险
10021	IP被封禁，请稍后再试或者联系客服
21000	训练任务内部异常，请重新提交或联系客服处理
21001	形象包训练任务提交异常
21002	找不到人像ID
21003	人像查询异常
50001	额度预扣除失败
50002	余额不足
其他	内部异常，请重新提交或联系客服处理

🔗 查询人像

🔗 接口说明

分页查询定制的数字人人像列表，接口鉴权和通用分页参数查阅：[接口通用说明](#)。

🔗 接口协议

GET /api/digitalhuman/open/v1/figure/lite2d/query

🔗 GET 请求参数

参数	类型	必填	说明
figureId	string	否	指定人像ID查询，为空则查询该账号下所有人像
systemFigure	boolean	否	是否查询平台的公共人像 1. 为空：查询全部 2. true：返回平台公共人像 3. false：返回定制人像
trainSuccess	boolean	否	是否查询训练完成，状态可用的人像 1. 为空：不进行过滤 2. true：只返回可用人像（平台公共人像+训练成功状态的定制人像） 3. false：只返回排队中、训练中或训练失败的定制人像
pageNo	int	否	页码，默认为1
pageSize	int	否	每页容量，默认10

🔗 返回参数

参数	类型	必填	说明
figureId	string	是	人像ID
name	string	是	数字人名称
customizeType	string	是	定制类型
systemFigure	boolean	是	是否是平台公共人像
gender	string	是	性别
taskSubmitTime	string	是	任务提交时间
templateImg	string	是	数字人缩略图
pictureUrl	string	是	参考图片地址（入参figureImageUrl对应的图片）
templateVideoUrl	string	是	训练视频URL
resolutionWidth	int	是	训练视频像宽
resolutionHeight	int	是	训练视频像高
status	string	是	状态 1. LINE_UP（排队中） 2. ACTIONING（动作生成中） 3. GENERATING（开始训练） 4. SUCCESS（成功） 5. FAILED（失败）
failedCode	int	否	失败错误码
failedMessage	string	否	制作失败原因

🔗 请求示例

```
GET /api/digitalhuman/open/v1/figure/lite2d/query?figureId=931
```

🔗 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "pageNo": 1,
    "pageSize": 100,
    "totalCount": 87,
    "result": [
      {
        "figureId": "267303",
        "name": "数字人test2025-dong12",
        "customizeType": "GENERATE_NO_ACTION",
        "systemFigure": false,
        "gender": "MALE",
        "templateVideoUrl": "https://meta-human-editor-test.bj.bcebos.com/vid-rfcsnigdkfrsrx8d.mp4",
        "resolutionWidth": 1080,
        "resolutionHeight": 1920,
        "status": "SUCCESS",
        "templateImg": "https://star-light-lite.cdn.bcebos.com/figure/video/fig-rfcrbxtg7ir6dura20250603_161448main.png",
        "taskSubmitTime": "2025-06-03 16:13:24",
        "failedCode": 0,
        "pictureUrl": "https://open-api-test.bj.bcebos.com/%E7%94%9F%E6%88%90%E5%BC%8F%E6%B5%8B%E8%AF%95%2FSlide%2016_9%20-%201.png"
      }
    ]
  },
  "requestId": "fd48ba29-12ec-47b4-a959-69db7a5a1cbc",
  "success": true
}
```

视频组件接入手册

123数字人视频

123数字人视频合成接口

🔗 接口说明

只需对着镜头说十秒123，即可快速输出口播成片

支持输入文本或音频驱动数字人生成视频，该功能仅支持固定背景视频

1、调用视频合成任务接口，提交数字人底板视频和驱动文本、音频，返回任务ID

2、接收任务结束的回调通知，或使用任务ID轮询任务状态（不建议）

接口调用域名：<https://open.xiling.baidu.com>

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

[123数字人拍摄指南](#)

🔗 接口列表

🔗 提交视频合成任务

接口说明 提交视频合成任务，返回任务ID和底板ID

接口协议 POST /api/digitalhuman/open/v1/video/submit/fast

Header Content-Type: application/json;charset=utf-8

请求参数 | 参数 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | templateVideoId | string | 是 | 底板视频素材文件 ID

素材文件ID支持两种方式：

1. 使用[文件上传接口](#)上传视频文件获取文件ID，其中 providerType 参数填写：OPEN_VIDEO_123_DURATION_VIDEO
2. 支持上传公网URL

视频要求：

1. 数量：1个
2. 大小：3GB以内
3. 时长：10秒 - 4分钟
4. 格式：mp4、mov
5. 分辨率：最大支持4K
7. 帧率：建议大于24fps（低帧率会转成25fps）
8. 色彩标准：SDR-Rec.709（非HDR） || driveType | string | 否 | 驱动数字人的数据类型，枚举值，默认 TEXT。
 - 1、TEXT：文本驱动，系统会调用 TTS 合成音频后驱动数字人
 - 2、VOICE：音频驱动，使用输入音频驱动数字人 || text | string | 否 | 驱动数字人播报的文本，当 driveType 为 TEXT 时必填
 - 1、字符长度不超过 20000
 - 2、支持 SSML 标签，使用说明见 [SSML使用说明](#) || ttsParams | object | 否 | TTS 参数，当 driveType 为 TEXT 时必填 || -- person | string | 否 | 发音人ID，可用发音人列表参考：[公共音色库](#) || -- lan | string | 否 | 注：仅有LITE版克隆音色及公共音色库中的支持多语言的音色支持该参数。
- 支持参数列表：'Chinese', 'Chinese,Yue', 'English', 'Russian', 'Spanish', 'French', 'Portuguese', 'German', 'Turkish', 'Dutch', 'Ukrainian', 'Vietnamese', 'Indonesian', 'Japanese', 'Italian', 'Korean', 'Thai', 'Polish', 'Romanian', 'Greek', 'Czech', 'Finnish', 'Hindi', 'auto' || -- speed | string | 否 | 语速：5是正常值，0-15的取值范围。越大语速越快，默认值5（字面量需要是整数） || -- volume | string | 否 | 音量：5是正常值，0-15的取值范围，越大音量越大，默认值5（字面量需要是整数） || -- pitch | string | 否 | 语调：5是正常值，0-15的取值范围，越大声音越尖，默认值5（字面量需要是整数） || inputAudioUrl | string | 否 | 驱动数字人播报的音频 url，url 字符长度不超过 1000。当 driveType 为 VOICE 时必填。音频要求：
 1. 时长不超过 90 分钟
 2. 文件大小不超过2G
 3. 支持的音频格式为wav、mp3、wma、m4a || callbackUrl | string | 否 | 接口调用方接受任务回调通知的url，规则和接口说明参考接口通用说明，回调消息定义如下表。 |

回调参数

- type: VIDEO_FAST_GENERATE
- data:

参数名	类型	是否必填	描述
taskId	string	是	任务ID
status	string	是	状态 1. WAIT（排队中） 2. GENERATING（开始训练） 3. SUCCESS（成功） 4. FAILED（失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
videoUrl	string	否	视频文件地址，文件会保存 7 天
duration	int	否	视频时长，单位 ms

返回参数 | 参数 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | taskId | string | 是 | 任务ID，用于调用查询接口查询任务状态 |

请求示例

```
{
  "templateVideoId": "30085e2d-b343-4b09-af25-a28430598230",
  "text": "有人说家是温馨的港湾。有人说家是心灵的寄托。如果您正在为家装而发愁，不妨考虑与我们合作。",
  "ttsParams": {
    "person": "20000000",
    "speed": "5",
    "volume": "5",
    "pitch": "5"
  },
}
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "vf3-rais2u9129541i0v"
  },
  "requestId": "bd15906f-ba35-4471-a3c2-b2cbf048189d",
  "success": true
}
```

错误码 | 错误码 | 描述 | | --- | --- | | 10006 | 参数校验不通过 | | 10011 | 没有购买视频时长或并发 | | 21001 | 视频任务提交异常 | | 21002 | 找不到底板ID | | 21003 | 数字人查询异常 | | 50001 | 额度预扣除失败 | | 50002 | 余额不足 |

🔍 查询任务

接口说明 使用任务ID查询任务状态

接口协议 GET /api/digitalhuman/open/v1/video/task

GET 请求参数 | 参数 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | taskId | string | 是 | 视频任务ID |

返回参数 | 参数 | 类型 | 必填 | 说明 | | --- | --- | --- | --- | | taskId | string | 否 | 视频任务ID | | videoUrl | string | 否 | 视频文件地址，文件会保存 7 天 | | duration | int | 否 | 视频时长，单位 ms | | createTime | string | 是 | 创建时间，示例：2024-04-29T20:28:06 | | updateTime | string | 是 | 更新时间，示例：2024-04-29T20:28:06 | | status | string | 是 | 状态

1. SUBMIT（已提交）
2. LINE_UP（排队中）
3. GENERATING（合成中）
4. SUCCESS（合成成功）
5. FAILED（合成失败） | | failedCode | int | 否 | 失败错误码 | | failedMessage | string | 否 | 制作失败原因 |

请求示例

```
GET /api/digitalhuman/open/v1/video/task?taskId=vf3-rajqt8jue9aiqu0p
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "vf3-rajqt8jue9aiqu0p",
    "status": "SUCCESS",
    "failedCode": 0,
    "failedMessage": "",
    "videoUrl": "https://digital-human-pipeline-output.cdn.bcebos.com/6780b9fde6929f677f3bb7cc_165.webm",
    "duration": 9167,
    "createTime": "2025-01-10T14:06:07",
    "updateTime": "2025-01-10T14:11:24",
    "startTrainTime": "2025-01-10T14:11:09+08:00"
  },
  "requestId": "d2a9e2c6-d751-4566-beb8-bda4bad97b7f",
  "success": true
}
```

123数字人拍摄指南

拍摄123数字人时，我们需要您提供一段**10秒-4分钟**口播123123的真人视频。为了提升最终生成的数字人效果，我们特别准备了以下拍摄指南，供您参考：

什么是123数字人？

- 只需对着镜头说十秒的123，就可以分钟级生成数字人口播成片

123数字人拍摄参考：

1、拍摄准备

1. 模特服饰
- 可按照模特喜好选择，与背景颜色风格协调、适合实际应用场景即可
 - 合身，不能过大或过紧，整洁大方，无明显褶皱

- 穿低领衣服，露出脖子
- 避免碎花、格纹、条纹等大面积花纹的衣服，避免摩尔纹产生

2. 模特妆造

- 妆容自然大方，有亲和力，无明显突兀感
- 如果存在瑕疵，例如皮肤不平整，长痘，手部存在血丝，需要遮瑕
- 露出双耳，可佩戴固定耳饰，不建议佩戴易晃动耳饰

3. 布景：

- 实景拍摄：背景整洁、无密集纹理，适合实际应用场景

4. 设备：

- 支持手机拍摄，有更高效果要求可使用相机
- 拍摄的分辨率不超过4K，关闭HDR模式，关闭log模式
- 比例：建议使用横屏为16:9，竖屏为9:16，效果最佳
- 角度：建议正对镜头，不要仰头、低头、仰视、俯视
- 构图：将主播放置于画面中间，确保五官清晰、面部无遮挡
- 可使用三脚架或稳定器固定拍摄设备，以减少画面抖动，确保视频流畅稳定

2、形象拍摄（重点关注）

1. 视频要求：**10秒-4分钟**，10秒生成会更快
2. 口播内容：只要说“**1231234567**”即可，可手持商品或者道具
3. 面部神情：头部保持相对稳定，面带微笑，保持放松
4. 模特动作：一些细微姿态动作会看起来更自然，更放松。**不建议剧烈动作和走动**

3、后期处理

1. 剪辑和调色美颜：

- 需掐头去尾剪辑，最终提供的123人像视频需全程无拼接，跳帧对视频效果影响较大
- 片段删减建议：可去掉开头和结尾未做好准备的部分
- 调色和美颜建议：如对效果有要求，可做轻微调色和美颜

2. 视频导出建议：

- 大小：3GB以内
- 时长：**10秒-4分钟**
- 格式：mp4、mov
- 分辨率：不超过4K
- 比例：建议16:9或9:16
- 帧率：25fps
- 色彩标准：SDR-Rec.709（非HDR）

照片数字人视频合成接口

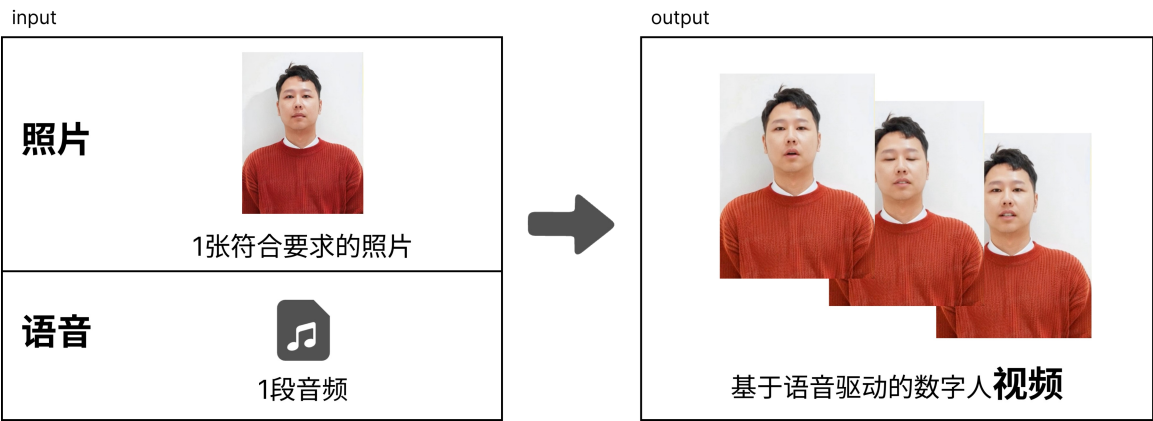
1. 接口说明

接口调用地址：<https://open.xiling.baidu.com>

语音驱动人像图片合成播报视频，支持合成透明背景视频。

- 调用视频合成任务接口，提交语音音频和图片，返回任务ID
- 接收任务结束的回调通知，或使用任务ID轮询任务状态（不推荐）

接口鉴权和通用字段说明请查阅：[接口通用说明](#)



注意：输入的照片需符合以下规范，否则输出视频会有异样



- 照片清晰，只包含单个人脸
- 面部正对镜头，轮廓完整
- 面部无明显阴影，无头发遮挡
- 无夸张表情，不露牙齿
- 背景无其他人脸、杂物
- 图片分辨率不超过1920*1920

2. 接口列表

2.1 提交视频合成任务

2.1.1 接口说明

提交语音音频和人像图片，返回任务ID。合成视频的分辨率和图片分辨率一致，若宽高超过1920*1920px，会自动等比缩小。

2.1.2 接口协议

POST /api/digitalhuman/open/v1/video/image/submit

Header Content-Type: application/json;charset=utf-8

2.1.3 请求参数

参数	类型	必填	说明
inputImageUrl	string	是	人像图片 url，url 字符长度不超过 1000，图片需要包含清晰的人脸，图片格式要求： 1.文件大小不超过 3 MB； 2.支持的图片格式png、jpg、jpeg；当图片是包含透明通道的 png 格式，会返回 webm 格式透明背景视频，其他情况返回 mp4 格式视频。 图片内容要求： 1.照片清晰，只有单个人脸； 2.面部正对镜头，轮廓完整； 3.面部无明显阴影，无头发遮挡； 4.无夸张表情，不露牙齿
driveType	string	否	驱动数字人的数据类型，枚举值，默认 TEXT。 1.TEXT：文本驱动，系统会调用 TTS 合成音频后驱动数字人。 2.VOICE: 音频驱动，使用输入音频驱动数字人。
text	string	否	驱动数字人播报的文本，当 driveType 为 TEXT 时必填，字符长度不超过 20000。输入文本需包含标点符号分割，连续文本不超过1024字节，每个中文一个字按3个字节，英文字母及数字空格按1个字节计算。支持 SSML 标签（仅限中文音色），使用说明见 SSML使用说明 • 当 lan 参数非'auto'时，请自行确认文本与所选语言是否一致，如不一致合成成功后会导致发音异常。
ttsParams	object	否	TTS 参数，当 driveType 为 TEXT 时必填
-- person	string	否	发音人ID，发音人列表详见 公共音色库
-- lan	string	否	注：仅有LITE版克隆音色及公共音色库中的支持多语言的音色支持该参数。 • 支持参数列表：'Chinese', 'Chinese,Yue', 'English', 'Russian', 'Spanish', 'French', 'Portuguese', 'German', 'Turkish', 'Dutch', 'Ukrainian', 'Vietnamese', 'Indonesian', 'Japanese', 'Italian', 'Korean', 'Thai', 'Polish', 'Romanian', 'Greek', 'Czech', 'Finnish', 'Hindi', 'auto'
-- speed	string	否	语速：5是正常值，0-15的取值范围。越大语速越快，默认值5（字面量需要是整数）
-- volume	string	否	音量：5是正常值，0-15的取值范围，越大音量越大，默认值5（字面量需要是整数）
-- pitch	string	否	语调：5是正常值，0-15的取值范围，越大声音越尖，默认值5（字面量需要是整数）
inputAudioUrl	string	否	驱动数字人播报的音频url，当 driveType 为 VOICE 时必填。url 字符长度不超过 1000，音频要求： 1.时长不超过 90 分钟且不小于2秒。 2.文件大小不超过 2G。 3.支持的音频格式为wav、mp3、wma、m4a。
subtitleParams	object	否	字幕参数，当 driveType 为 TEXT 时生效
-- subtitlePolicy	string	否	字幕策略，枚举值，默认：SRT，字幕信息会以 SRT 格式保存到字幕文件中，在视频信息可以获取文件 URL，字段名：subtitleFileUrl
-- enabled	boolean	否	是否开启，默认：false
callbackUrl	string	否	接口调用方接受任务回调通知的url，规则和接口说明参考 接口通用说明 ，回调消息定义如下表。

2.1.4 回调参数

- type: VIDEO_GENERATE_IMAGE
- data:

参数名	类型	是否必填	描述
taskId	string	是	任务ID
status	string	是	状态： SUCCESS（合成成功） FAILED（合成失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
videoUrl	string	否	视频文件地址，文件会保存 7 天
duration	int	否	视频时长，单位 ms
subtitleFileUrl	string	否	字幕文件地址，文件会保存 7 天

2.1.5 返回参数

参数	类型	必填	说明
taskId	string	是	任务ID，用于调用任务查询接口查询任务状态

2.1.6 请求示例

```
{
  "requestId": "xxx11",
  "text": "嗨！我是挥手问候家，为你打造专属问候，不论早晚，挥手間温暖送达！",
  "driveType": "TEXT",
  "ttsParams": {
    "person": 5116,
    "speed": 5,
    "pitch": 5,
    "volume": 5
  },
  "subtitleParams": {
    "enabled": false,
    "subtitlePolicy": "SRT"
  },
  "inputImageUrl": "https://meta-human-editor-prd.cdn.bcebos.com/ca1d70b6f73d44bebd2aba1f3deb60bb/draftThumbnail/dt-qfdkzu6nhawzayqb5.png",
  "callbackUrl": "http://api-gateway:8080/api/digitalhuman/internal/api-gateway/v1/callback/vp/test"
}
```

2.1.7 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "img-qgti819vzahqxp3",
    "status": null,
    "failedCode": 0,
    "failedMessage": null,
    "videoUrl": null,
    "duration": 0,
    "createTime": null,
    "updateTime": null,
    "subtitleFileUrl": null
  },
  "requestId": "xxxx11",
  "success": true
}
```

2.1.8 错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
14001	内部服务异常，请稍后再试
14002	内部服务异常，网络拥堵
20001	未检测到合格人脸
20004	检测到多个人脸
20002	inputImageUrl 文件下载失败
20003	inputAudioUrl 文件下载失败
22006	人像提交未知异常
50001	额度预扣除失败
50003	额度扣除失败
50004	余额不足

2.2 查询任务

2.2.1 接口说明

使用任务ID查询任务状态

2.2.2 接口协议

GET /api/digitalhuman/open/v1/video/image/task

2.2.3 GET 请求参数

参数	类型	必填	说明
taskId	string	是	任务ID

2.2.4 返回参数

参数	类型	必填	说明
taskId	string	是	任务ID
status	string	是	状态： SUBMIT（已提交待合成） GENERATING（合成中） SUCCESS（合成成功） FAILED（合成失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
videoUrl	string	否	视频文件地址，文件会保存 7 天
duration	int	否	视频时长，单位 ms
subtitleFileUrl	string	否	字幕文件 URL
createTime	string	是	创建时间，示例：2024-04-29T20:28:06
updateTime	string	是	更新时间，示例：2024-04-29T20:28:06

2.2.5 请求示例

```
GET
/api/digitalhuman/open/v1/video/image/task?taskId=abc123&requestId=123e4567-e89b-12d3-a456-426614174000
```

2.2.6 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "img-qgti819vzahqxp3",
    "status": "GENERATING",
    "failedCode": 0,
    "failedMessage": "",
    "videoUrl": null,
    "duration": 0,
    "createTime": "2024-07-18T08:17:56",
    "updateTime": "2024-07-18T08:17:56",
    "subtitleFileUrl": null
  },
  "requestId": "fdc66116-762d-4a82-a082-e83357a3167e",
  "success": true
}
```

2.3 人像图片校验

2.3.1 接口说明

检测人像图片是否符合视频合成的要求。非必须，如果业务场景交互流程需要可以提前做人像校验。

2.3.2 接口协议

POST /api/digitalhuman/open/v1/video/image/verify

Header Content-Type: application/json;charset=utf-8

2.3.3 请求参数

参数	类型	必填	说明
inputImageUrl	string	是	人像图片 url，需要包含清晰的人脸，图片格式要求： 1.文件大小不超过 3 MB； 2.支持的图片格式为png、jpg、jpeg。 图片内容要求： 1.照片清晰，只有单个人脸； 2.面部正对镜头，轮廓完整； 3.面部无明显阴影，无头发遮挡； 4.无夸张表情，不露牙齿

2.3.4 返回参数

无，未通过时返回错误码和错误信息。

2.3.5 请求示例

```
{
  "requestId": "123e4567-e89b-12d3-a456-426614174000",
  "inputImageUrl": "https://example.com/image.png"
}
```

2.3.6 返回示例

```
{
  "requestId": "123e4567-e89b-12d3-a456-426614174000",
  "code": 0,
  "success": true,
  "message": {
    "global": "success"
  }
}
```

2.3.7 错误码

错误码	描述
20001	未检测到合格人脸
20004	检测到多个人脸

基础视频合成接口

1. 接口说明

接口调用地址：<https://open.xiling.baidu.com>

输入文本或音频驱动数字人生成播报视频，支持合成透明背景视频。

- 1. 调用视频合成任务接口，提交数字人配置和驱动文本、音频，返回任务ID
- 2. 接收任务结束的回调通知，或使用任务ID轮询任务状态（不建议）

接口鉴权和通用字段说明请查阅[接口通用说明](#)

2. 接口列表

2.1 提交视频合成任务

2.1.1 接口说明

配置数字人播报文本或音频，以及其他视频参数，提交合成任务，返回任务ID。

2.1.2 接口协议

POST /api/digitalhuman/open/v1/video/submit

Header Content-Type: application/json;charset=utf-8

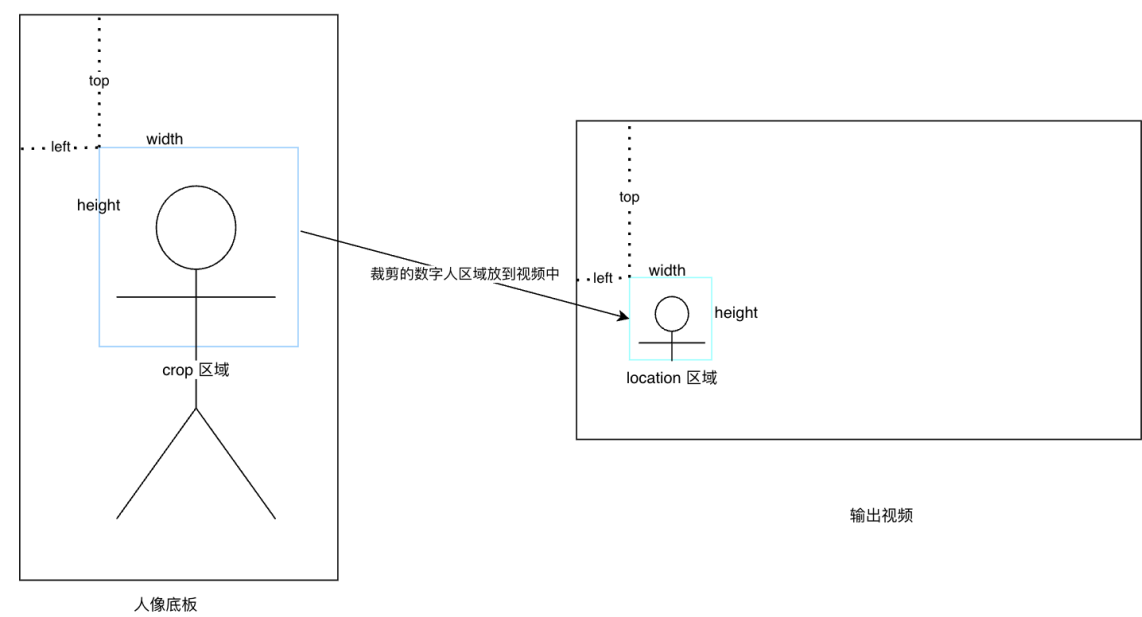
2.1.3 请求参数

参数	类型	必填	说明
figureId	string	是	人像 ID，参考人像说明文档： 3D数字人 、 2D精品数字人 、 2D小样本数字人
driveType	string	否	驱动数字人的数据类型，枚举值，默认 TEXT。 1.TEXT：文本驱动，系统会调用 TTS 合成音频后驱动数字人 2.VOICE: 音频驱动，使用输入音频驱动数字人
text	string	否	驱动数字人播报的文本，当 driveType 为 TEXT 时必填，字符长度不超过 20000。输入文本需包含标点符号分割，连续文本不超过1024字节，每个中文一个字按3个字节，英文字母及数字空格按1个字节计算。支持 SSML 标签，使用说明见[SSML使用说明]（仅限中文音色） https://cloud.baidu.com/doc/AI_DH/s/rlyyduk53 • 当 lan 参数非'auto'时，请自行确认文本与所选语言是否一致，如不一致合成成功会导致发音异常。
ttsPara	object	否	TTS 参数，当 driveType 为 TEXT 时必填

ms	object	否	无参数，当 driveType 为 TEXT 时必填
-- person	string	否	发音人ID，可用发音人列表参考： 公共音色库 ；如果开通的声音克隆服务，也可以使用用户自己克隆的音色ID作为发音人
-- lan	string	否	注：仅有LITE版克隆音色及公共音色库中的支持多语言的音色支持该参数。 • 支持参数列表：'Chinese', 'Chinese,Yue', 'English', 'Russian', 'Spanish', 'French', 'Portuguese', 'German', 'Turkish', 'Dutch', 'Ukrainian', 'Vietnamese', 'Indonesian', 'Japanese', 'Italian', 'Korean', 'Thai', 'Polish', 'Romanian', 'Greek', 'Czech', 'Finnish', 'Hindi', 'auto'
-- speed	string	否	语速：5是正常值，0-15的取值范围。越大语速越快，默认值5（字面量需要是整数）
-- volume	string	否	音量：5是正常值，0-15的取值范围，越大音量越大，默认值5（字面量需要是整数）
-- pitch	string	否	语调：5是正常值，0-15的取值范围，越大声音越尖，默认值5（字面量需要是整数）
inputAudioUrl	string	否	驱动数字人播报的音频 url，url 字符长度不超过 1000。当 driveType 为 VOICE 时必填。音频要求： 1.时长不超过 90 分钟 2.文件大小不超过2G 3.支持的音频格式为wav、mp3、wma、m4a
playControlParams	object	否	控制合成视频中人像底板视频的播放逻辑，不传则使用底板视频从头开始正反循环播放。只有用户定制的2d小样本人像可以使用此功能，公共人像，精品人像不能使用。
-- controlType	String	是	控制类型，枚举值： 1. SEGMENT_LOOP：指定底板视频片段正反循环播放，通过startTimestamp和endTimestamp选取片段；
-- startTimestamp	double	否	起始时间戳，单位是秒，支持两位小数，起始时间要小于底板视频总时长，不填默认取 0。
-- endTimestamp	double	否	结束时间戳，单位是秒，支持两位小数，结束时间不能大于底板视频总时长且要大于起始时间（起始时间和结束时间的间隔要大于0.04秒，也就是1帧），不填默认取视频结尾。
videoParams	object	是	视频参数
-- width	int	是	视频分辨率，最大不超过 4k，并且不能超过人像底板宽或高的分辨率
-- height	int	是	
-- transparent	boolean	否	是否输出 webm 格式透明背景视频，false 则输出mp4格式视频，默认 true
dhParams	object	否	数字人参数，控制数字人的大小位置，目前支持三种方式，生效优先级 positionV2 > position > camerald
-- cameraId	int	否	数字人机位，预置了数字人的位置和大小，支持的人像类型： 3D数字人 、 2D精品数字人 ，机位ID 点击查阅人像介绍文档
-- position	object	否	数字人位置和大小，不能和 camera id 同时指定，支持的人像类型： 3D数字人 、 2D精品数字人 、 2D小样本数字人
-- v	float	否	水平位置调整数，范围 -0.5 到 0.5，取 0 居中，填充数字人填充左

-- -- x	float	否	水平位置调整，范围 -0.5 到 0.5，取 0 居中，越大数字人越靠右
-- -- y	float	否	竖直位置调整，范围 -1 到 1，取 0 数字人底部对齐视频底部，越大数字人越靠上
-- -- z	float	否	缩放比例调整，范围 0.5 到 1.5，取1原始比例，越小数字人越小。
-- positionV2	object	否	通过剪裁方式像素级别控制数字人位置和大小，如果启用此参数，dhParams中的position 和 cameraId 参数无效，目前支持 2D小样本数字人 。 先根据剪裁参数（crop）从人像底板选取数字人矩形区域，然后使用位置参数（location）控制数字人区域在最终视频里的位置和大小，参考表格下面的示例图。
-- -- crop	object	否	裁剪参数，默认不裁剪，即取整个人像底板
-- -- -- top	int	否	数字人区域到人像底板顶边距离，单位：像素，默认取 0 表示取顶边，不能超过人像底板的高度
-- -- -- left	int	否	数字人区域到人像底板左边距离，单位：像素，默认取 0 表示取左边，不能超过人像底板的宽度
-- -- -- width	int	是	数字人区域的宽度，单位：像素，不能小于底板宽度的 0.1 倍，left + width 不能超过底板的宽度
-- -- -- height	int	是	数字人区域的宽度，单位：像素，不能小于底板高度的 0.1 倍，top + height 不能超过底板的高度
-- -- location	object	是	数字人区域在视频里的位置
-- -- -- top	int	否	数字人区域到视频顶边距离，单位：像素，默认取 0 表示和顶边重合，不能超过视频的高度
-- -- -- left	int	否	数字人区域到视频左边距离，单位：像素，默认取 0 表示和左边重合，不能超过视频的宽度
-- -- -- width	int	是	视频中数字人区域的宽度，单位：像素，不能随意缩放，取值范围：[crop.width * 0.3, crop.width * 1.5]
-- -- -- height	int	是	视频中数字人区域的高度，单位：像素，不能随意缩放，取值范围：[crop.height * 0.3, crop.height * 1.5]
subtitleParams	object	否	字幕参数，当 driveType 为 TEXT 时生效
-- subtitlePolicy	string	否	字幕策略，枚举值，默认：SRT，字幕信息会以 SRT 格式保存到字幕文件中，在视频信息可以获取文件 URL，字段名：subtitleFileUrl
-- enable	boolean	否	是否开启，默认：false
backgroundImageUrl	string	否	背景图片 url，字符长度不超过 1000，图片长宽比例建议和视频分辨率保持一致，否则会进行拉伸或填充黑边。图片格式要求： 1.文件大小不超过 3 MB 2.支持的图片格式：png、jpg、jpeg、bmp
callbackUrl	string	否	接口调用方接受任务回调通知的url，规则和接口说明参考 接口通用说明 ，回调消息定义如下表。
autoAnimoji	boolean	否	是否自动添加数字人动作，默认值 false。生效条件： 1.值设为 true 2.人像是3D 数字人2D 精品数字人 3.驱动方式是 TEXT

positionV2 参数 crop 及 location 示例图：



2.1.4 回调参数

- type: VIDEO_GENERATE
- data:

参数名	类型	是否必填	描述
taskId	string	是	任务ID
status	string	是	状态： SUCCESS（合成成功） FAILED（合成失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
videoUrl	string	否	视频文件地址，文件会保存 7 天
duration	int	否	视频时长，单位 ms
subtitleFileUrl	string	否	字幕文件地址，文件会保存 7 天

2.1.5 返回参数

参数	类型	必填	说明
taskId	string	是	任务ID，用于调用任务查询接口查询任务状态

2.1.6 请求示例

```
{
  "figureId": "1081",
  "text": "嗨！我是挥手问候家，为你打造专属问候，不论早晚，挥手间温暖送达！",
  "driveType": "TEXT",
  "backgroundImageUrl": "https://meta-human-editor-
prd.cdn.bcebos.com/ca1d70b6f73d44bebd2aba1f3deb60bb/draftThumbnail/dt-qfdkzu6nhawzayqb5.png",
  "ttsParams": {
    "person": "5116",
    "speed": "5",
    "pitch": "5",
    "volume": "5"
  },
  "videoParams": {
    "height": 1920,
    "width": 1080,
    "transparent": false
  },
  "autoAnimoji": false,
  "subtitleParams": {
    "enabled": false,
    "subtitlePolicy": "SRT"
  },
  "callbackUrl": "http://api-gateway:8080/api/digitalhuman/internal/api-gateway/v1/callback/vp/test"
}
```

2.1.7 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "vid-qgti66nzkch2xy9h",
    "status": null,
    "failedCode": 0,
    "failedMessage": null,
    "videoUrl": null,
    "duration": 0,
    "createTime": null,
    "updateTime": null,
    "subtitleFileUrl": null
  },
  "requestId": "b978a051-fafe-4623-bfa5-1c07566e88ac",
  "success": true
}
```

2.1.8 错误码

错误码	描述
20003	inputAudioUrl 文件下载失败
20005	backgroundImageUrl 文件下载失败
20034	图片下载异常
20005	backgroundImageUrl 文件下载失败
50001	额度预扣除失败

2.2 查询任务

2.2.1 接口说明

使用任务ID查询任务状态

2.2.2 接口协议

GET /api/digitalhuman/open/v1/video/task

2.2.3 GET 请求参数

参数	类型	必填	说明
taskId	string	是	任务ID

2.2.4 返回参数

参数	类型	必填	说明
taskId	string	是	任务ID
status	string	是	状态： SUBMIT（已提交待合成） GENERATING（合成中） SUCCESS（合成成功） FAILED（合成失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
videoUrl	string	否	视频文件地址，文件会保存 7 天
duration	int	否	视频时长，单位 ms
subtitleFileUrl	string	否	字幕文件 URL
createTime	string	是	创建时间，示例：2024-04-29T20:28:06
updateTime	string	是	更新时间，示例：2024-04-29T20:28:06

2.2.5 请求示例

```
GET /api/digitalhuman/open/v1/video/task?taskId=abc123&requestId=123e4567-e89b-12d3-a456-426614174000
```

2.2.6 返回示例

```
{
  "requestId": "123e4567-e89b-12d3-a456-426614174000",
  "code": 0,
  "success": true,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "abc123",
    "status": "SUCCESS",
    "failedCode": null,
    "failedMessage": null,
    "videoUrl": "https://example.com/video.mp4",
    "duration": 300000,
    "createTime": "2024-04-29T20:28:06",
    "updateTime": "2024-04-29T20:28:06"
  }
}
```

高级视频合成接口

1. 接口说明

接口调用地址：<https://open.xiling.baidu.com>

基于预置模板样式，传入播报文本、视频标题、logo、背景图等素材，合成数字人播报视频。

- 调用视频合成任务接口，提交数字人配置和驱动文本、返回任务ID
- 接收任务结束的回调通知，或使用任务ID轮询任务状态（不推荐）

接口鉴权和通用字段说明请查阅[接口通用说明](#)

2. 接口列表

2.1 提交视频合成任务

2.1.1 接口说明

配置数字人播报文本或音频，以及其他视频参数，提交合成任务，返回任务ID。

2.1.2 接口协议

POST /api/digitalhuman/open/v1/video/advanced/submit

Header Content-Type: application/json;charset=utf-8

2.1.3 请求参数 | 参数 | 类型 | 必填 | 说明 || :---: | :---: | :---: | --- || figureId | string | 是 | 人像 ID，人像列表详见：[2D小样本数字人](#)（仅支持2D小样本**站姿**人像） || templateId | string | 是 | 模板 ID，详见 附.《高级视频模板列表》 注：使用模版时，暂不支持使用多语言功能 || driveType | string | 否 | 驱动数字人的数据类型，枚举值，默认 TEXT。 1.TEXT：文本驱动，系统会调用 TTS 合成音频后驱动数字人 2.VOICE: 音频驱动，使用输入音频驱动数字人 || inputAudioUrl | string | 否 | 驱动数字人播报的音频 url，url 字符长度不超过 1000。当 driveType 为 VOICE 时必填。

音频要求：

- 时长不超过 10 分钟
 - 文件大小不超过120 MB
 - 支持的音频格式为wav、mp3、wma、m4a || text | string | 是 | 驱动数字人播报的文本，字符长度不超过 2000。输入文本需包含标点符号分割，连续文本不超过1024字节，每个中文一个字按3个字节，英文字母及数字空格按1个字节计算。支持 SSML 标签（仅限中文音色），使用说明详见：[SSML使用说明](#)
- 当 lan 参数非'auto'时，请自行确认文本与所选语言是否一致，如不一致合成成功后会致发音异常。 || title | string | 否 | 视

频标题，标题的文本内容，如果缺省则视频中没有标题，长度限制30 || logoParams | object | 否 | logo | | -- logoUrl | string | 否 | logo图片的 url，如果缺省则视频中没有logo，长度限制1000，

图片格式要求：

1. 文件大小不超过 3 MB

2. 支持的图片格式：png、jpg、jpeg、bmp || bgmParams | object | 否 | bgm | | -- bgmUrl | string | 否 | bgm 的 url，长度限制1000，

音频要求：

1. 文件大小不超过 120 MB

2. 支持的音频格式：wav、mp3、wma、m4a || materialUrl | string | 否 | 背景素材 url，支持图片和视频，长度限制1000，视频长度大于播报长度截取前面的部分，小于则循环播放。

图片格式要求：

1. 文件大小不超过 3 MB

2. 支持的图片格式：png、jpg、jpeg、bmp

视频格式要求：

1. 文件大小不超过 500 MB

2. 支持的视频格式：mp4 || ttsParams | object | 是 | TTS 参数 || -- person | string | 否 | 发音人ID，音色库列表详见：[公共音色库](#) || -- lan | string | 否 | 注：仅有LITE版克隆音色及公共音色库中的支持多语言的音色支持该参数。•支持参数列

表：'Chinese', 'Chinese,Yue', 'English', 'Russian', 'Spanish', 'French', 'Portuguese', 'German', 'Turkish', 'Dutch', 'Ukrainian', 'Vietnamese', 'Indonesian', 'Japanese', 'Italian', 'Korean', 'Thai', 'Polish', 'Romanian', 'Greek', 'Czech', 'Finnish', 'Hindi',

'auto' || -- speed | string | 否 | 语速：5是正常值，0-15的取值范围。越大语速越快，默认值5（字面量需要是整数） || -- volume | string | 否 | 音量：5是正常值，0-15的取值范围，越大音量越大，默认值5（字面量需要是整数） || -- pitch | string | 否 | 语调：5是正常值，0-15的取值范围，越大声音越尖，默认值5（字面量需要是整数） || videoParams | object | 是 | 视频

参数 || -- width | int | 是 | 视频分辨率，需要和模板分辨率比例保持一致，最大不超过 1080x1920 或 1920x1080 || -- height | int | 是 | 视频分辨率，需要和模板分辨率比例保持一致，最大不超过 1080x1920 或 1920x1080 || callbackUrl | string | 否 | 接口调用方接受任务回调通知的 url，规则和接口说明参考[接口通用说明](#) 回调消息定义如下表 || riskTip | string | 否 | 水印描述，

不超过30个字，默认底部居中（原saas水印字段，已有，这次开放） || openingMaterial | object | 否 | 片头背景素材，传一个素材对象，格式：1. 视频格式：mp4 2. 文件大小不超过 500 MB || -- fileId | string | 否 | 使用openingMaterial时候，fileId和

fileUrl二者必须选一必填，fileId，使用[文件上传接口](#)接口上传后获取文件ID，providerType 参数固定填写 OPEN_VIDEO_DURATION_VIDEO，如果fileId和fileUrl都填写，会用fileId，fileUrl无效 || -- fileUrl | string | 否 | 使用 openingMaterial时候，fileId和fileUrl二者必须选一必填，文件的url（必须是bos地址，bcebos.com域名下，url地址中不要携带

<, >, ", #, %, {, }, |, ^, ~, [,], 以及空格等非法字符） || -- mediaType | string | 否 | 使用openingMaterial必填，媒体类型，枚举：VIDEO | endingMaterial | object | 否 | 片尾背景素材，传一个素材对象，格式：1. 视频格式：mp4 2. 文件大小不超过 500 MB | -- fileId | string | 否 | 使用endingMaterial时候，fileId和fileUrl二者必须选一必填，fileId，使用[文件上传接口](#)

(<http://doc.bce.baidu-int.com/@bce-doc/file/document/8a7e969b-a045-4999-a49d-9817e9488e75>)接口上传后获取文件ID，providerType 参数固定填写 OPEN_VIDEO_DURATION_VIDEO，如果fileId和fileUrl都填写，会用fileId，fileUrl无效 || -- fileUrl | string | 否 | 使用endingMaterial时候，fileId和fileUrl二者必须选一必填，文件的url（必须是bos地址，bcebos.com域名下，url地址中不要携带

<, >, ", #, %, {, }, |, ^, ~, [,], 以及空格等非法字符） || -- mediaType | string | 否 | 使用 endingMaterial必填，媒体类型，枚举：VIDEO || mashupMaterials | object | 否 | 混剪背景素材，传一个素材对象的数组，最多可以传10个，视频切割成5-8s，图片随机播放3-5s，把所有素材混合后随机播放，格式：1. 视频格式：mp4 2. 图片格式：png、jpg、jpeg、bmp 3. 文件大小视频不超过 500 MB，图片不超过3MB || -- fileId | string | 否 | 使用mashupMaterials时候，fileId和fileUrl二者必须选一必填，fileId，使用[文件上传接口](#)接口上传后获取文件ID，providerType 参数固定填写

OPEN_VIDEO_DURATION_VIDEO，如果fileId和fileUrl都填写，会用fileId，fileUrl无效 || -- fileUrl | string | 否 | 使用 mashupMaterials时候，fileId和fileUrl二者必须选一必填，文件的url（必须是bos地址，bcebos.com域名下，url地址中不要携带

<, >, ", #, %, {, }, |, ^, ~, [,], 以及空格等非法字符） || -- mediaType | string | 否 | 使用mashupMaterials必填，媒体类型，枚举：VIDEO，IMAGE || fissionParams | object | 否 | 裂变参数，使用figureId和ttsPerson组合生成多个视频 || -- figureIds | List | 否 | 人像裂变参数，使用fissionParams必传，人像id数组，传入会使用裂变功能，会覆盖figureId参数，最多可以传10个人像ID || -- ttsPersons | List | 否 | 音色列表参数，音色id数组，如果如果不传默认只按人像id裂变，如果传入且driveType为text时会覆盖ttsParam的person参数，如果driveType为voice时，该tts参数无效，只按figureId进行裂变，最多可以传10个音色值。

注：多语言功能暂不支持在裂变中使用。 || extendBackgroundUrl | string | 否 | 扩展自定义背景，填充视频外的背景

支持填写素材 url，支持图片和视频，长度限制1000，视频长度大于播报长度截取前面的部分，小于则循环播放。

图片格式要求：

- 1.文件大小不超过3 MB
- 2.支持的图片格式：png、jpg、jpeg、bmp

视频格式要求：

- 1.文件大小不超过500 MB
- 2.支持的视频格式：mp4 ||

2.1.4 回调参数

- type: VIDEO_GENERATE_ADVANCED
- data:

参数名	类型	是否必填	描述
taskId	string	是	任务ID
status	string	是	状态： SUCCESS（合成成功） FAILED（合成失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
videoUrl	string	否	视频文件地址，文件会保存 7 天
duration	int	否	视频时长，单位 ms

2.1.5 返回参数

参数	类型	必填	说明
taskId	string	是	任务ID，用于调用任务查询接口查询任务状态
fissionTasks	List<object>	否	使用裂变功能（请求的fissionParams有值且通过校验生效）时返回，格式为一个对象数组，包含任务执行的信息
-- taskId	string	否	任务id，使用是原参数的taskId无效，code为0是才返回的才是有效的id
-- message	string	否	任务执行的具体情况，成功为SUCCESS
-- code	int	否	错误码，0为正常返回
-- figureId	string	否	人像id，当前task使用的人像id
-- ttsPerson	string	否	音色值，当前task使用的音色值

2.1.6 请求示例

```

{
  "figureId": "1081",
  "templateId": "t-ad4eeqsspfzwyqxyte125",
  "text": "有人说家是温馨的港湾。有人说家是心灵的寄托。如果您正在为家装而发愁，不妨考虑与我们合作。让我们为您解决装修的烦恼，让您的家焕然一新，焕发新生。点击下方联系我们，获取福利吧。",
  "inputAudioUrl": "https://digital-human-pipeline-output.cdn.bcebos.com/vis/audio/a-qewefy6hd5mwaxik4s.wav",
  "driveType": "TEXT",
  "title": "家装盛典福利等你来",
  "logoParams": {
    "logoUrl": "https://meta-human-editor-test.cdn.bcebos.com/7e430c5bd2084a00b90a83f5190f6fec/f47fc33a-d63f-4bd0-815b-387cd0f373fc/logo%E7%99%BD.png"
  },
  "ttsParams": {
    "person": "5116",
    "speed": "5",
    "volume": "5",
    "pitch": "5"
  },
  "riskTip": "@baidu水印",
  "openingMaterial": {
    {
      "fileUrl": "https://meta-human-editor-test.bj.bcebos.com/fashion-test/output.mp4",
      "mediaType": "VIDEO"
    }
  },
  "endingMaterial": {
    {
      "fileUrl": "https://meta-human-editor-test.bj.bcebos.com/fashion-test/output2.mp4",
      "mediaType": "VIDEO"
    }
  },
  "mashupMaterials": [
    {
      {
        "fileUrl": "https://meta-human-editor-test.bj.bcebos.com/fashion-test/0fea7f8524371b55eaf5eb987.jpg",
        "mediaType": "IMAGE"
      }
    },
    {
      {
        "fileUrl": "https://meta-human-editor-test.bj.bcebos.com/fashion-test/7052b61826fae54158fea02ba.png",
        "mediaType": "IMAGE"
      }
    },
    {
      {
        "fileUrl": "https://meta-human-editor-test.bj.bcebos.com/fashion-test/output3.mp4",
        "mediaType": "VIDEO"
      }
    }
  ],
  "videoParams": {
    "width": 1080,
    "height": 1920
  },
  "fissionParams": {
    "figureIds": ["1081",
      "1112"],
    "ttsPersons": ["4100",
      "5132"]
  }
}

```

新增请求示例（扩展自定义背景示例）

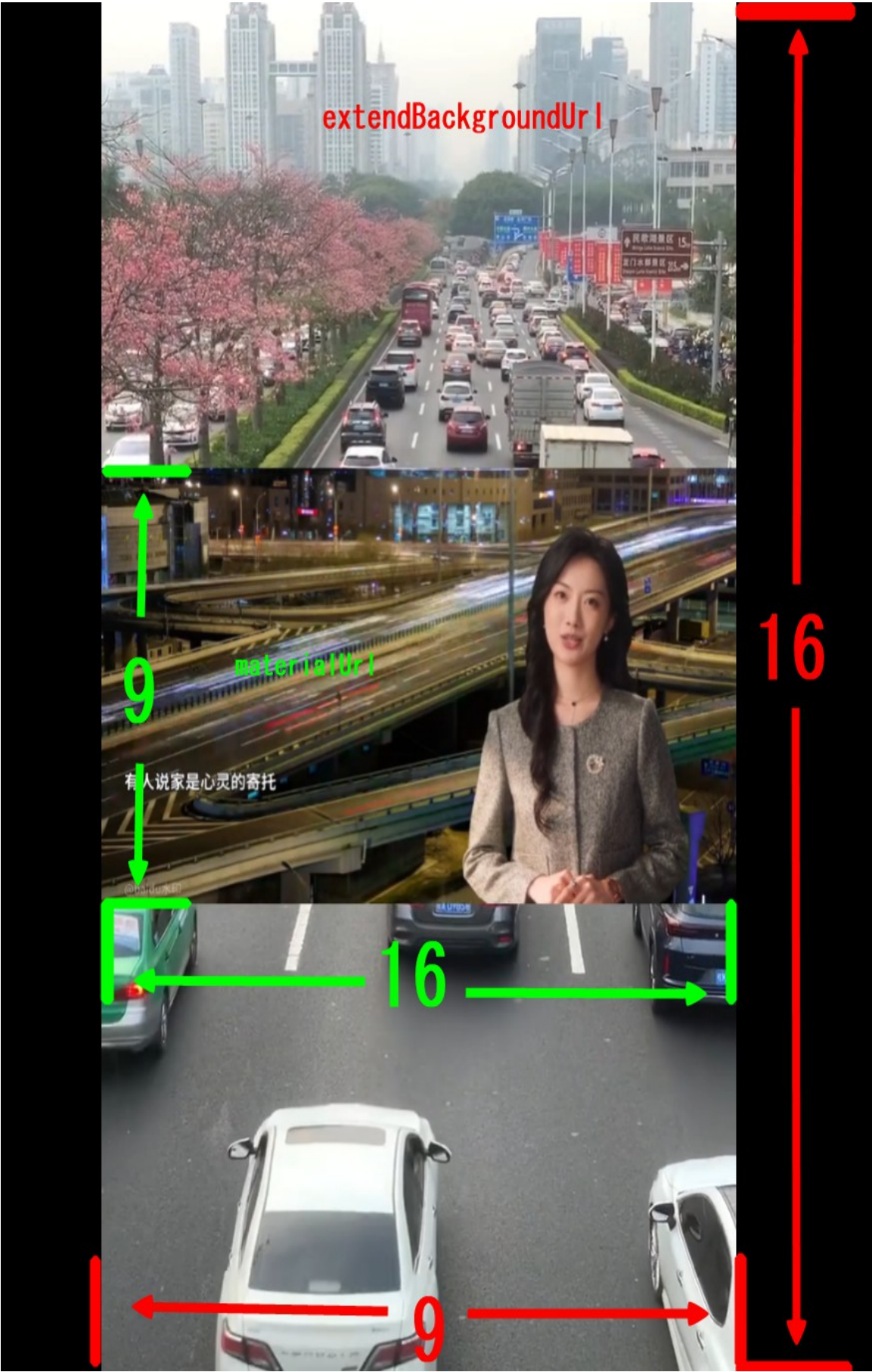
使用 extendBackgroundUrl 参数，请注意templateId和materialUrl参数，这两个参数的比例与视频相反。

- 制作横屏视频：视频分辨率(videoParams)和扩展自定义背景(extendBackgroundUrl)尺寸比例为 16:9。而模板（templateId）和背景（materialUrl）需要使用比例为 9:16 的素材

- 制作竖屏视频：视频分辨率(videoParams)和扩展自定义背景(extendBackgroundUrl)尺寸比例为 9:16。而模板 (templateId) 和背景 (materialUrl) 需要使用比例为 16:9 的素材
- 视频的尺寸会根据背景大小等比缩放，比如扩展背景尺寸1280x720, 则视频及背景 (materialUrl) 将自动缩放为406x720

```
{
  "figureId": "211868",
  "templateId": "t-cd4eeqsspfzwyqxyteditu4",
  "text": "有人说家是温馨的港湾。有人说家是心灵的寄托。如果您正在为家装而发愁，不妨考虑与我们合作。让我们为您解决装修的烦恼，让您的家焕然一新，焕发新生。点击下方联系我们，获取福利吧。",
  "materialUrl": "https://meta-human-editor-test.bj.bcebos.com/bg_720_1280.mp4",
  "extendBackgroundUrl": "https://meta-human-editor-test.bj.bcebos.com/bg_1280_720.mp4",
  "driveType": "TEXT",
  "ttsParams": {
    "person": "5116",
    "speed": "5",
    "volume": "5",
    "pitch": "5"
  },
  "riskTip": "@baidu水印",
  "videoParams": {
    "width": 1280,
    "height": 720
  }
}
```

参考图示：



2.1.1.7 返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": null,
    "status": null,
    "failedCode": 0,
    "failedMessage": null,
    "videoUrl": null,
    "duration": 0,
    "createTime": null,
    "updateTime": null,
    "fissionTasks": [
      {
        "taskId": "adv-qmutfwsptw57zegc-1112-4100",
        "message": "SUCCESS",
        "figureId": "xxx",
        "ttsPerson": "xxx",
        "code": 0
      },
      {
        "taskId": "adv-qmutfwsptw57zegc-1081-5132",
        "message": "SUCCESS",
        "figureId": "xxx",
        "ttsPerson": "xxx",
        "code": 0
      },
      {
        "taskId": "adv-qmutfwsptw57zegc-1081-4100",
        "message": "SUCCESS",
        "figureId": "xxx",
        "ttsPerson": "xxx",
        "code": 0
      },
      {
        "taskId": "adv-qmutfwsptw57zegc-1112-5132",
        "message": "SUCCESS",
        "figureId": "xxx",
        "ttsPerson": "xxx",
        "code": 0
      }
    ]
  },
  "requestId": "c5434cbd-c144-4a5a-8133-d6ef46c5f84b",
  "success": true
}
```

2.1.8 错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
14001	内部服务异常，请稍后再试
14002	内部服务异常，网络拥堵
50001	冻结额度失败
20005	背景图下载失败
20036	视频任务提交失败
10006	参数校验失败

2.2查询任务接口

2.2.1 接口说明

使用任务ID查询任务状态

2.2.2 接口协议

GET /api/digitalhuman/open/v1/video/advanced/task

2.2.3 GET 请求参数

参数	类型	必填	说明
taskId	string	是	任务ID

2.2.4 返回参数

参数	类型	必填	说明
taskId	string	是	任务ID
status	string	是	状态： READY_FOR_SUBMIT（收到请求，处理请求参数，准备提交） SUBMIT（已提交待合成） GENERATING（合成中） SUCCESS（合成成功） FAILED（合成失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
videoUrl	string	否	视频文件地址，文件会保存 7 天
duration	int	否	视频时长，单位 ms
createTime	string	是	创建时间，示例：2024-04-29T20:28:06
updateTime	string	是	更新时间，示例：2024-04-29T20:28:06

2.2.5 请求示例

```
GET
/api/digitalhuman/open/v1/video/advanced/task?taskId=adv-qgti5xyxx6e8yxu&requestId=123e4567-e89b-12d3-a456-42e612345678
0
```

2.2.6 返回示例

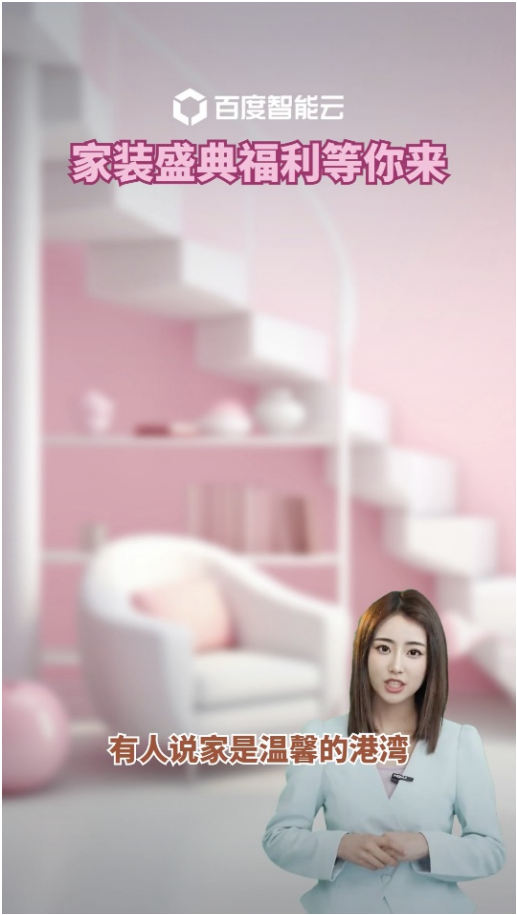
```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "adv-qhfrtb627i772vvj",
    "status": "GENERATING",
    "failedCode": 0,
    "failedMessage": "OK",
    "videoUrl": "",
    "duration": 20060,
    "createTime": "2024-08-06T15:17:46",
    "updateTime": "2024-08-06T15:17:46",
    "subtitleFileUrl": null
  },
  "requestId": "f5d816ec-1938-4fd7-960c-f140b65414d2",
  "success": true
}
```

附. 高级视频模板列表

序号	templateId	比例	模板预览	demo视频
				

1	t-pf4kqasspwzwyexyte121	竖屏（9:16）		1
2	t-af4keqsspzfzwyexyte123	竖屏（9:16）		
				

3	t-ad4eeqsspzfzwyqxyte125	竖屏（9:16）		
4	t-cd4eeqsspzfzwyqxyte127	竖屏（9:16）		
5	t-hb4eeqsspzfzwyqxyte129	横屏（16:9）		
				

6	t-cd4eeqsspzfzwyqxyteditu3	竖屏（9:16）	
7	t-cd4eeqsspzfzwyqxyteditu4	竖屏（9:16）	
8	t-cd4eeqsspzfzwyqxyteditu5 (无数字人模板)	竖屏（9:16）	



语音合成接口

接口说明

新增功能:支持多语言音频合成

接口调用域名：<https://open.xiling.baidu.com>

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

异步任务方式合成语音音频，同步接口和流式接口暂不开放。音频合成需要购买【交互组件-语音合成】组件，操作参考：[平台操作指引](#)

接口列表

提交音频合成任务

POST /api/digitalhuman/open/v1/tts/text2audio/submit

Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必填	说明
text	string	是	字符长度不超过 2000。支持 SSML 标签（仅限中文音色），使用说明见 SSML使用说明 • 当 lan 参数非'auto'时，请自行确认文本与所选语言是否一致，如不一致合成成功后会导致发音异常。
person	string	是	发音人ID，可用发音人列表参考： 公共音色库 ，或克隆自有音色： 音色克隆和查询接口
lan	string	否	注：仅有LITE版克隆音色及公共音色库中的支持多语言的音色支持该参数。 • 支持参数列表：'Chinese', 'Chinese,Yue', 'English', 'Russian', 'Spanish', 'French', 'Portuguese', 'German', 'Turkish', 'Dutch', 'Ukrainian', 'Vietnamese', 'Indonesian', 'Japanese', 'Italian', 'Korean', 'Thai', 'Polish', 'Romanian', 'Greek', 'Czech', 'Finnish', 'Hindi', 'auto'
speed	int	否	语速：5是正常值，0-15的取值范围。越大语速越快，默认值5
volume	int	否	音量：5是正常值，0-15的取值范围，越大音量越大，默认值5
pitch	int	否	音调：5是正常值，0-15的取值范围。越大音调越高，默认值5
outputFormat	string	否	音频输出格式，支持mp3和wav，默认mp3
sampleRate	int	否	输出音频的每秒采样次数，目前只支持16000
callbackUrl	string	否	接口调用方接受任务回调通知的url，规则和接口说明参考 任务回调说明 ，回调消息字段见下面回调参数表格。

返回参数

参数	类型	说明
taskId	string	任务ID，用于调用任务查询接口查询任务状态

错误码

错误码	描述
23002	音色ID不存在
10006	参数校验不通过
50004	余额不足

请求示例

```
{
  "text": "今天阳光好好呀，适合出去玩",
  "person": "5132",
  "speed": 5,
  "volume": 5,
  "pitch": 5,
  "outputFormat": "mp3",
  "sampleRate": 16000
}
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "t2a-xxx"
  },
  "requestId": "a829316c-aa14-42be-839e-d2e432b58cca",
  "success": true
}
```

回调参数

- type: AUDIO_GENERATE
- data:

参数名	类型	是否必填	描述
taskId	string	是	任务ID
status	string	是	状态： SUCCESS（合成成功） FAILED（合成失败）
failedCode	int	否	失败错误码
failedMessage	string	否	失败错误信息
audioUrl	string	否	音频文件地址，文件会保存 7 天
duration	int	否	音频时长，单位 ms
textTimestamp	Array<Sentence>	是	文本时间戳信息
-- content	string	是	拆分的句子内容
-- startTimestamp	long	是	句子开始时间戳，单位毫秒（ms）
-- endTimestamp	long	是	句子结束时间戳，单位毫秒（ms）

🔗 查询音频合成任务状态

GET /api/digitalhuman/open/v1/tts/text2audio/task

🔗 请求参数

参数	类型	必填	说明
taskId	string	是	任务ID

🔗 返回参数

参数	类型	必填	说明
status	string	是	状态： SUBMIT（已提交待合成） GENERATING（合成中） SUCCESS（合成成功） FAILED（合成失败）
failedCode	int	否	失败错误码，具体见下面表格
failedMessage	string	否	失败错误信息
audioUrl	string	否	音频文件地址，文件会保存 7 天
duration	int	否	音频时长，单位 ms
textTimestamp	Array<Sentence>	是	文本时间戳信息
-- content	string	是	拆分的句子内容
-- startTimestamp	long	是	句子开始时间戳，单位毫秒（ms）
-- endTimestamp	long	是	句子结束时间戳，单位毫秒（ms）
createTime	string	是	创建时间，示例：2024-04-29T20:28:06
updateTime	string	是	更新时间，示例：2024-04-29T20:28:06

failedCode 说明

错误码	描述
24001	语音合成服务调用异常
24002	音频编码失败
24010	语音合成任务执行失败

🔗 请求示例

GET /api/digitalhuman/open/v1/tts/text2audio/task?taskId=t2a-xxx

🔗 返回示例

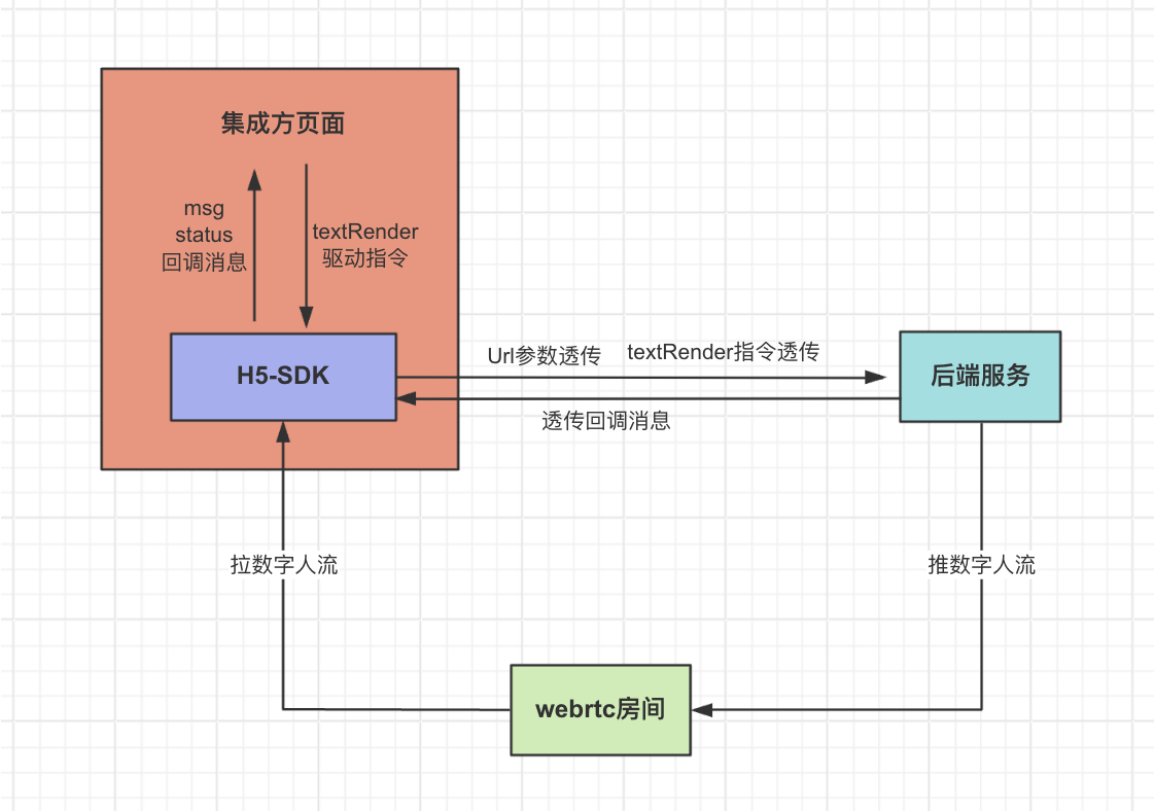
```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "taskId": "t2a-xxx",
    "status": "SUCCESS",
    "audioUrl": "https://meta-human-editor-test.bj.bcebos.com/xxxx.mp3",
    "duration": 3110,
    "textTimestamp": [
      {
        "content": "今天阳光好好呀",
        "startTimestamp": 0,
        "endTimestamp": 1720
      },
      {
        "content": "适合出去玩",
        "startTimestamp": 1720,
        "endTimestamp": 3110
      }
    ]
  },
  "createTime": "2025-03-26T14:27:39",
  "updateTime": "2025-03-26T14:28:12"
},
"requestId": "010c6b17-7fcd-4a61-9602-e8eece2602e",
"success": true
}
```

交互组件接入手册

云渲染交互

H5版

🔗 1. SDK运行原理



2. 接入方法

iframe支持在网页内引入其他网页，使用iframe引入WebSdk方式如下：

FAQ: 为什么使用iframe

使用iframe的方式，可以天然的和集成方代码形成隔离，形成单独的进程进行数字人的渲染工作，防范因js单线程设计可能造成的渲染卡顿问题。

```
// 用户侧
<iframe
  id="myIframe"
  src="https://open.xiling.baidu.com/cloud/realtime?token=自行填写&initMode=noAudio&camerald=0"
  frameborder="0"
  allow="microphone;camera;midi;encrypted-media;autoplay;" // 可用权限
/>
```

3. demo使用

3.1 demo说明

此demo分别使用react，vue两种框架。针对如何发送消息驱动数字人，如何监听驱动数字人消息的完成状态，如何第一时间判断数字人已加载完成

如何解决浏览器静音策略，以及如何节省数字人资源等问题，进行了代码实现。

3.2 demo下载

React版本，下载地址<https://sdk-demo.bj.bcebos.com/realtime-digital-human-demo.zip>

Vue版本，下载地址<https://sdk-demo.bj.bcebos.com/realtime-digital-human-vue-demo.zip>

核心代码如下

```

/* eslint-disable @typescript-eslint/no-explicit-any */
// 参考文献 https://developer.mozilla.org/en-US/docs/Web/Media/Autoplay_guide
import { useCallback, useEffect, useState } from 'react'
import {v4 as uuidV4} from 'uuid';
import DHIframe from '@bddh/starling-dhiframe';

import useAudioRender from './hook/useAudioRender';

import './App.css'

const dhiframe = new DHIframe('digital-human-iframe');

enum ReadyState {
  UNINSTANTIATED = -1,
  CONNECTING = 0,
  OPEN = 1,
  CLOSING = 2,
  CLOSED = 3,
}

// 可以有声播放则返回true,否则返回false
function checkPlayUnMute(): Promise<boolean> | boolean {
  if (typeof navigator.userActivation !== 'undefined' && typeof navigator.userActivation.hasBeenActive !== 'undefined') {
    return navigator.userActivation.hasBeenActive;
  }
  return new Promise<(resolve) => {
    const audioElem: HTMLAudioElement = document.createElement('audio');
    audioElem.src =
'data:audio/wav;base64,UklGRigAAABXQVZFZm10IBAAAAABAAEASsAACJWAAACABAAZGF0YQAAAAA=';
    audioElem.muted = false;

    const playPromise: Promise<void> | undefined = audioElem.play();

    if (playPromise !== undefined) {
      playPromise
        .then(() => {
          // 自动播放成功
          resolve(true);
        })
        .catch((error: Error) => {
          if (error.name === "NotAllowedError" || error.name === "AbortError") {
            // 自动播放被禁止或中止
            resolve(false);
          } else {
            // 其他错误
            resolve(false);
          }
        });
      audioElem.remove();
    } else {
      // 如果 play() 返回 undefined，可能是浏览器不支持 Promise 风格的 play()
      audioElem.remove();
      resolve(false);
    }
  });
}

function App() {
  // 实时视频流远端准备完毕，可以进行textRender等驱动指令的发送
  const [realTimeVideoReady, setRealTimeVideoReady] = useState(false);
  const [wsConnected, setWsConnected] = useState(false);
  // 浏览器限制导致的有声视频无法自动播放，所以这里进行静音后自动播放

```

```

const [videolsMuted, setVideolsMuted] = useState(false);
const [commandId] = useState(uuidV4());
// 是否可以自动播放，检测完成状态
const [checkOver, setCheckOver] = useState(false);
// 超时即将消失tip
const [showTimeoutTip, setShowTimeoutTip] = useState(false);
// 数字人前置loading页面
const [showLoadingPage, setShowLoadingPage] = useState(false);

const [playAudio, playMultiAudio] = useAudioRender(dhlframe); // pcm格式音频播放

const onMessage = useCallback((msg: any) => {
  if (msg.origin === 'https://open.xiling.baidu.com') {
    const {type, content} = msg.data;
    const {action, requestId} = content;
    switch (type) {
      case 'rtcState':
        if (content.action === 'remoteVideoConnected') {
          setRealTimeVideoReady(true);
        }
        if (content.action === 'localVideoMuted' && content.body) {
          setVideolsMuted(true);
        }
        break;
      case 'wsState':
        if (content.readyState === ReadyState.OPEN) {
          setWsConnected(true);
        }
        else if (content.readyState === ReadyState.CLOSED || content.readyState === ReadyState.CLOSING) {
          setWsConnected(false);
        }
        break;
      case 'msg':
        if (requestId === commandId && action === 'FINISHED') {
          console.info(`数字人驱动完成, 驱动id为${requestId}`);
          const newCommandId = uuidV4();
          // 解开注释可以进行持续播报
          // setCommandId(newCommandId);
          dhlframe.sendMessage({
            action: 'TEXT_RENDER',
            body: '我已经完成一次驱动，这是我说的第n句话',
            requestId: newCommandId
          });
        }
        else if (action === 'DISCONNECT_ALERT') {
          setShowTimeoutTip(true);
        }
        else if (action === 'TIMEOUT_EXIT') {
          const iframeDom = document.getElementById('digital-human-iframe');
          iframeDom?.remove();
        }
        break;
      default:
        break;
    }
  }
}, [commandId]);

// 播放开场白
const playWelcome = () => {
  setVideolsMuted(false);
  dhlframe.sendCommand({

```

```

        subType: 'muteAudio',
        subContent: false
    });
}

// 打断数字人
const handleInterrupt = () => {
    dhIframe.sendMessage({
        action: 'TEXT_RENDER',
        body: '<interrupt></interrupt>',
        requestId: commandId
    });
}

const goNextPage = () => {
    setShowLoadingPage(false);
    setVideolsMuted(false);
}

useEffect(() => {
    const checkPlayUnMuteFun = async () => {
        const result = await checkPlayUnMute(); // 500ms左右
        setVideolsMuted(!result);
        setCheckOver(true);
    }
    checkPlayUnMuteFun();
}, [])

useEffect(() => {
    if (realTimeVideoReady && wsConnected && checkOver && !videolsMuted) {
        dhIframe.sendMessage({
            action: 'TEXT_RENDER',
            body: '这是我的开场白自我介绍，我是数字人',
            requestId: commandId
        });
    }
}, [realTimeVideoReady, wsConnected, checkOver, videolsMuted, commandId]);

useEffect(() => {
    dhIframe.registerMessageReceived(onMessage);
    return () => {
        dhIframe.removeMessageReceived(onMessage);
    };
}, [onMessage]);

if (showLoadingPage) {
    return (
        <div onClick={goNextPage}>
            前置loading页，点击进入下个页面
        </div>
    )
}

return (
    <>
        {videolsMuted && <div className="tip1" onClick={playWelcome}>取消静音，播报开场白</div>}
        {showTimeoutTip && <div className="tip2" onClick={playWelcome}>长时间无交互，数字人马上消失了，点我一下，进行交互</div>}
        <div className="button-wrapper">
            <button onClick={playWelcome}>文本播报</button>
            <button onClick={playAudio}>音频播报</button>
        </div>
    </>
)

```

```
    <button onClick={playMultiAudio}>多段拼接音频播报</button>
    <button onClick={handleInterrupt}>打断播报</button>
  </div>
  <iframe
    id="digital-human-iframe"
    src="https://open.xiling.baidu.com/cloud/realtime?token=自行填写&initMode=noAudio&cameraId=自行填写
    &figureId=自行填写"
    width="450"
    height="800"
    allow="autoplay"
  >
  </iframe>
</>
)
}
```

export default App

3.3 demo运行和修改

1. 解压后，进入目录：React版本目录名称：realtime-digital-human-demo、Vue 版本目录名称：realtime-digital-human-vue-demo
2. 判断环境是否有 node(22 版本以上)和 yarn，如没有则进行 node, yarn 的安装：brew install node; npm install --global yarn (其他安装方式请自行搜索)
3. 命令行执行 yarn，安装所需依赖包
4. 命令行执行 npm run dev，启动项目
5. 将控制台输出的ip地址，复制到浏览器地址栏进行项目预览，页面黑色并显示在 loading 中
6. 在平台购买云渲染交互组件[平台操作指南](#)，新建应用关联组件，获取应用的appld和appkey
7. 在[公共形象库](#)中查找组件对应的人像类型的人像 id 和机位 id，替换 demo 文件中 figureId 和 cameraId 参数，文件位置：React 版本是 src/App.tsx、Vue 版本是 src/App.vue，搜索『需替换』关键字
8. 根据[接口通用说明](#)生成鉴权参数 token，替换 demo 文件中 token 参数，文件位置：React 版本是 src/App.tsx、Vue 版本是 src/App.vue，搜索『自行填写』关键字
9. 保存文件，刷新页面可以看到数字人成功加载出来，可以点击播报按钮驱动数字人播报

🔗 4. url参数说明

4.1 常用基础参数

参数	类型	必填	默认	备注
token	string	是	-	平台购买所需组件 平台操作指南 ,生成鉴权凭证:生成方法查阅 接口通用说明 (token=Authorization)的鉴权参数生成部分
figureId	string	是	-	查找文档 公共形象库 ，填写对应 figureId
initMode	String	是	-	用于针对特定应用场景快速批量设置关联配置的默认值，设置后关联配置无需填写，枚举类型： noAudio ：无拾音模式
resolutionWidth	string		当前设备宽度 clientWidth	视频分辨率宽，只能是偶数，最小值400，分辨率最大不超过 1080x1920、1920x1080
resolutionHeight	string		当前设备高度 clientHeight	视频分辨率高，只能是偶数，最小值400，分辨率最大不超过 1080x1920、1920x1080
cameralId	int		-	数字人预设机位 ID，控制数字人的位置和大小，建议使用和视频分辨率比例一致的机位：高精 2D 人像：0（横屏半身）、1（竖屏半身）详细枚举参考 3D数字人 、 2D精品数字人

4.2 高级参数

参数	类型	默认	备注
mode	string	inline	画面裁剪参数（inline内嵌展示未填充的地方背景色展示、crop裁剪）
backgroundImageUrl	string	-	设置数字人video背景图片，不支持直接配置颜色 autoChromaKey为 true 时不生效
videoBg	string	-	设置页面颜色，默认黑色，iframe 宽高比和数字人视频宽高比不一致时，页面的填充色， 大多情况建议使用 transparent autoChromaKey为 true 时不生效
autoChromaKey	Boolean	false	开启自动抠绿，实现数字人透明背景
chromaEffects	Object	-	自定义抠绿参数，使用时需将autoChromaKey设置为false;详解设置步骤参考: 常见问题7.11
ttsPer	string	-	音色发音人ID，可用发音人列表参考： 公共音色库
cp-ttsLan	string	-	注：仅有LITE版克隆音色及公共音色库中的支持多语言的音色支持该参数。 • 支持参数列表：'Chinese', 'Chinese,Yue', 'English', 'Russian', 'Spanish', 'French', 'Portuguese', 'German', 'Turkish', 'Dutch', 'Ukrainian', 'Vietnamese', 'Indonesian', 'Japanese', 'Italian', 'Korean', 'Thai', 'Polish', 'Romanian', 'Greek', 'Czech', 'Finnish', 'Hindi', 'auto' •泰语需要使用标点进行句子分隔
ttsPitch	number	5	音调，5是正常值，0-15的取值范围，越大声音越尖，默认值5（字面量需要是整数）
ttsSpeed	number	5	语速，5是正常值，0-15的取值范围。越大语速越快，默认值5（字面量需要是整数）
ttsVolume	number	5	音量，5是正常值，0-15的取值范围，越大音量越大，默认值5（字面量需要是整数）
cp-position	Object	-	通过剪裁方式像素级别控制数字人位置和大小，目前支持: 先根据剪裁参数（crop）从人像底板选取数字人矩形区域，然后使用位置参数（location）控制数

nV2			字人区域在最终视频里的位置和大小。支持通过可视化操作生成该参数： 人像配置台
-- crop	Object	-	裁剪参数，默认不裁剪，即取整个人像底板
-- -- top	number	-	数字人区域到人像底板顶边距离，单位：像素，默认取 0 表示取顶边，不能超过人像底板的高度
-- -- left	number	-	数字人区域到人像底板左边距离，单位：像素，默认取 0 表示取左边，不能超过人像底板的宽度
-- -- width	number	-	数字人区域的宽度，单位：像素，不能小于底板宽度的 0.1 倍，left + width 不能超过底板的宽度
-- -- height	number	-	数字人区域的宽度，单位：像素，不能小于底板高度的 0.1 倍，top + height 不能超过底板的高度
-- location	Object	-	数字人区域在视频里的位置
-- -- top	number	-	数字人区域到视频顶边距离，单位：像素，默认取 0 表示和顶边重合，不能超过视频的高度
-- -- left	number	-	数字人区域到视频左边距离，单位：像素，默认取 0 表示和左边重合，不能超过视频的宽度
-- -- width	number	-	视频中数字人区域的宽度，单位：像素，不能随意缩放，取值范围：[crop.width * 0.3, crop.width * 1.5]
-- -- height	number	-	视频中数字人区域的高度，单位：像素，不能随意缩放，取值范围：[crop.height * 0.3, crop.height * 1.5]
cp-autoAni moji	Boolean	false	是否开启自动添加数字人动作，只支持高精3D/2D人像
cp- inactive DisconnectSec	int	180	长时间无交互（交互指发送 text render 让数字人播报）的断连提醒，单位秒，0 表示不开启。达到阈值时，服务端会发送 TIMEOUT_EXIT 提醒消息，父页面在收到该消息时需销毁子页面，避免数字人渲染资源浪费，目前数字人后端服务不会主动断开。
cp- preAlert Sec	int	10	在开启长时间无交互断连提醒，且即将到达断连阈值时，后端会先发送一条 DISCONNECT_ALERT 预提醒消息，父页面收到消息可以弹窗提醒用户。该参数用于设置断连前多少秒发送该提醒，比如无交互超时参数 cp-inactiveDisconnectSec 设置为 180（3 分钟），提醒参数 cp-preAlertSec 设置为 10（秒），则在用户最后一次交互后的 2 分 50 秒时会收到后端发送的 DISCONNECT_ALERT 提醒消息，在 3 分钟时收到 TIMEOUT_EXIT 断连消息。
showLogo	Boolean	false	是否展示logo
logoUrl	string	-	自定义logo图片地址
entry	Boolean	false	是否有入口页（开启后不会自动播放） 一般用于手机端iframe集成，手机端无法自动播放

4.3 debug参数设置

参数	类型	默认	备注
appId	string	-	应用标识和密钥，用于鉴权，前端直接使用可能会有泄漏风险，不建议生产环境使用，推荐使用 token 鉴权。
appKey	string	-	
showMessage	Boolean	false	展示toast提示消息
textAssist	Boolean	false	是否展示文本辅助, 可以进行驱动
showDebugger	Boolean	false	展示vconsole
closeLog	Boolean	false	关闭sdk侧控制台日志

🔗 5. 如何与Web-Sdk通信

使用iframe的内外层window可以通过[window.postMessage](#)这个API发送消息，基于[window.postMessage](#)封装好的通信npm包@bddh/starling-dhiframe，提供了收、发消息的方法，用户可直接使用。

5.1 @bddh/starling-dhiframe的安装和使用 工具包的安装

```
// npm 方式 :
npm i @bddh/starling-dhiframe
// yarn 方式
yarn add @bddh/starling-dhiframe
```

工具包的使用

```
import DHiframe from '@bddh/starling-dhiframe';
// digital-human-iframe需要和 iframe 的 id 地址保持一致
const dhiframe = new DHiframe('digital-human-iframe');
// 使用方式
dhiframe.sendMessage({
  action: 'TEXT_RENDER',
  body: '你好，这是我的开场白自我介绍，我是数字人',
  requestId: commandId
});
```

5.2 接收消息

用户侧可以通过如下方式接收leafletWeb抛出的消息

```
const onMessage = msg => {
  console.log(msg.data);
  // {type, content} = msg.data;
};

// 增加监听
dhiframe.registerMessageReceived(onMessage);
// 移除监听
dhiframe.removeMessageReceived(onMessage);
```

5.2.1 消息Type类型

用户侧能够接收到的消息（即Web-sdk抛出的消息）有三类，通过msg.data的type字段表示，三类消息说明如下：

type	content	备注
msg	msg	数字人驱动消息回调通知
rtcState	rtcState	webrtc 的状态
wsState	wsState	webSocket 的连接状态

5.2.2 rtcState数据信息

type为 rtcState 的 content 的 格式 {action, body}

action	body	备注
success	true	rtc房间登录成功
error	error	rtc房间登录失败
remotevideoloading	id	远端视频流开始加载
remoteVideoConnected	id	成功获取远端视频流（不受视频暂停影响）
remotevideoon	id	远端视频流加载完毕
localVideoMuted	true	首次加载无交互导致的video静音

5.2.3 wsState数据信息

type为 wsState 的 content.readyState

状态	value	备注
UNINSTANTIATED	-1	默认值 无状态
CONNECTING	0	ws 连接中
OPEN	1	ws 开启成功
CLOSING	2	ws 关闭中
CLOSED	3	ws 关闭

```
enum ReadyState {
  UNINSTANTIATED = -1,
  CONNECTING = 0,
  OPEN = 1,
  CLOSING = 2,
  CLOSED = 3,
}
```

5.2.4 msg数据信息

type为 msg 的 content.action

- 类别：驱动数字人回调消息

状态	备注	内容数据
RENDER_START	数字人收到驱动后，驱动开始	-
FINISHED	数字人完成驱动指令，驱动完成	-
RENDER_ERROR	文本驱动错误	-
DOWN_SUBTITLE	播报内容	body
RENDER_INTERRUPTED	数字人驱动被打断	-

- 类别：数字人退出回调消息

状态	备注	内容数据
TIMEOUT_EXIT	超时退出消息	

- 类别：数字人初始化消息回调

状态	备注	内容数据
CONNECT	连接消息	body
DISCONNECT_ALERT	长时间无交互提示	

type为 msg 的 **content.code**

- 错误类型：connect连接，初始化数字人报错报错

错误码	错误信息	说明
1001	服务器内部错误	
1002	APP已失效	
1003	APP token已过期	
1004	目前线上用户较多，请您稍后再试	app 会话路数已经达到上限
1005	授权信息不存在	figureId未授权不可用
1009	App key非法	
1011	裁剪参数非法	仅2D数字人存在
1012	参数不能为空	
1014	分辨率/编码参数不合法	
1015	无效的媒体信息	例如：传递的RTMP参数无效
1202	正在连接或已连接	
1204	解析请求失败	
1206	读取json失败	
2001	Fail to establish	建立渲染任务失败
3001	目前线上用户较多，请您稍后再试	系统渲染资源不足导致

- 错误类型：驱动数字人报错

错误码	错误信息	说明
-1	body非法	body内容含有非法字符
1013	DRML非法	仅2D数字人存在，不支持做该动作同时做播报，通常是走入走出的情况，面部不完整
1208	富文本格式非法	DRML格式错误
1201	操作尚未支持	

5.3 发送消息

用户侧一般需要发送的消息类型有两类。

一类是发给sdk对应后端服务的消息，如驱动或查询数字人；

另一个是指令型消息，如音频静音、视频播放等。

5.3.1 sendMessage方式

playVideo针对不同消息类型分别封装了不同的方法，可直接使用。

代码实例

```
import {v4 as uuidV4} from 'uuid';
// data 格式类型
const jsonStr = {action: 'TEXT_RENDER', body: '你好', requestId: uuidV4()}; // requestId不加默认使用sdk自生成的id
// 发送消息
dhlframe.sendMessage({
  action: 'TEXT_RENDER',
  body: '你好，这是我的开场白自我介绍，我是数字人',
  requestId: commandId
});
```

requestId说明：可自定义构造requestId进行驱动数字人，驱动的回调消息会沿用自定义构造的requestId。

action说明：可选值：TEXT_RENDER，含义为进行数字人语音驱动

body说明：

TEXT_RENDER时body的可选值	备注
你好，我是数字人	播报内容 • 当 cp-ttsLan 参数非'auto'时，请自行确认文本与所选语言是否一致，如不一致合成成功后会导致发音异常。
<interrupt></interrupt>	打断当前数字人播报
<say-as type="telephone">110</say-as></say-as>	指定读法的高级播报内容，支持 SSML使用说明
<silence time="5s"></silence>停止5s 后继续说话</silence>	支持播报中停顿
<audio src="http://meitron-test.bj.bcebos.com/mofachengbao_part.wav" /></audio>	音频播放
<interruptible="false">这段话不允许打断这段话不允许打断这段话不允许打断</interruptible>	播报不受打断的内容，后续textRender会在上句播报完成后开始播报，但仍可被interrupt标签打断

驱动消息回调事件，查看 [【5.2.4 msg数据信息】](#)

5.3.2 流式文本驱动方式

代码实例

```
dhIframe.sendMessage({
  action: 'TEXT_STREAM_RENDER',
  body: JSON.stringify({
    first: true,
    last: false,
    text: '你好啊，我是'
  }),
  requestId: commandId
});

dhIframe.sendMessage({
  action: 'TEXT_STREAM_RENDER',
  body: JSON.stringify({
    first: false,
    last: false,
    text: '数字人'
  }),
  requestId: commandId
});

dhIframe.sendMessage({
  action: 'TEXT_STREAM_RENDER',
  body: JSON.stringify({
    first: false,
    last: false,
    text: '分身'
  }),
  requestId: commandId
});

setTimeout(() => {
  dhIframe.sendMessage({
    action: 'TEXT_STREAM_RENDER',
    body: JSON.stringify({
      first: false,
      last: false,
      text: '今天天气'
    }),
    requestId: commandId
  });

  dhIframe.sendMessage({
    action: 'TEXT_STREAM_RENDER',
    body: JSON.stringify({
      first: false,
      last: true,
      text: '很好'
    }),
    requestId: commandId
  });
}, 500);
```

参数说明

参数	类型	解释
action	TEXT_STREAM_RENDER	流式文本驱动key
body	string	body: JSON.stringify({ text: string, (播报内容) first: boolean, (第一个流式数据设置first为true, 其余设置为false) last: boolean (最后一个流式数据设置last为true, 其余设置为false) })
requestId	string	生成的 随机id ，一轮流式文本驱动需使用同一个requestId

5.3.3 sendCommand方式

实例代码

```
// 驱使数字人视频静音指令
dhIframe.sendCommand({
  subType: 'muteAudio',
  subContent: mute // mute: true/false
});
```

subType参数

subType可选值	说明	dhIframe中已封装好的方法	使用方式
muteAudio	静音本地vidio, audio元素	muteAudio	dhIframe.muteAudio(true)
playVideo	数字人视频被暂停时播放	playVideo	dhIframe.playVideo(true)

5.3.4 sendAudioData方式

实例代码：流式音频驱动

```
// 第一条驱动音频
const requestId = uuidV4();
dhlframe.sendAudioData({
  action: 'AUDIO_STREAM_RENDER',
  requestId,
  body: JSON.stringify({
    audio: base64String, // 大小不可超过64k
    first: true,
    last: false
  })
})

// 第二条驱动音频
dhlframe.sendAudioData({
  action: 'AUDIO_STREAM_RENDER',
  requestId,
  body: JSON.stringify({
    audio: base64String,
    first: false,
    last: false
  })
})

// 第三条驱动音频
dhlframe.sendAudioData({
  action: 'AUDIO_STREAM_RENDER',
  requestId,
  body: JSON.stringify({
    audio: base64String,
    first: false,
    last: true
  })
})
```

tip1 : 仅支持pcm数据, pcm数据大小 限制**48K**

tip2 : 音频属性仅支持**"16k"**、**"16bit"** 和 **"单声道"**

16k : 这通常指的是音频的采样率为16kHz（千赫兹），意味着音频信号每秒被采样16,000次。采样率越高，能够记录的声音频率范围越宽，声音的高频细节就越丰富。16kHz的采样率一般适用于语音录音，足以覆盖人声的频率范围。

16bit : 这表示音频的位深度是16位。位深度决定了音频波形的动态范围，即最小声音和最大声音之间的区别。16位音频可以提供96分贝（dB）的动态范围，适合大多数音乐和语音应用。

单声道：单声道（Mono）音频指的是所有声音都混合到一个单一的音频通道中。与立体声（Stereo）不同，立体声至少使用两个通道（通常是左右两个通道），提供空间感和方向感。单声道适合许多应用，尤其是在只需要语音清晰传达的场合，比如新闻报道、有声书等。

参数解析

参数	类型	解释
action	AUDIO_STREAM_RENDER	AUDIO_STREAM_RENDER: 流式音频驱动播报
body	string	流式音频驱动 pcm数据大小不超过48K, audioBase64大小不超过64K body: JSON.stringify({ audio: base64String, (注: 将pcm音频数据转换为base64格式) first: boolean, (第一个流式数据设置first为true, 其余设置为false) last: boolean (最后一个流式数据设置last为true, 其余设置为false))})
requestId	string	生成的随机id

🔗 6. sdk兼容性说明

1、pc端-web测试结论

- windows系统
- 系统版本号：版本：windows 10 专业版 版本号：22H2 操作系统内部版本：19045.4170

说明：--表示无法安装对应版本的浏览器

浏览器	最新版本	拉流	最低兼容版本	拉流
Chrome	129.0.6668.101	✓	107.0.5304.107	✓
Safari	--	--	--	--
Edge	129.0.2792.89	✓	118.0.2088.69	✓
QQ浏览器	13.1.0	✓	12.1.1	✓

- MAC系统
- 系统版本号 芯片：Apple M1 序列号：C02F149YQ05D macOS:13.6（22G120）（Edge最低兼容版除外）
- 备注：Edge mac最低兼容使用系统版本：

芯片：Intel 序列号：FVFC226ZL412 macOS：13.6.4（22G513）

浏览器	最新版本	拉流	最低兼容版本	拉流
Chrome	129.0.6668.101	✓	107.0.5304.110	✓
Safari	16.6	✓	16.6	✓
Edge	129.0.2792.89	✓	107.0.1418.28	✓
QQ浏览器	5.0.7.203	✓	--	--

2、移动端-web测试结论

- Android

说明：--表示无法安装对应的浏览器

手机型号	系统	微信内置	QQ浏览器	Chrome	内置浏览器	UC浏览器
华为Mate 50	鸿蒙4.2.0	✓	✓	✓	✓	×
华为Mate 9	鸿蒙2.0.0	✓	✓	--	✓	×
vivo IQOOneo5	安卓13	✓	✓	✓	✓	×
vivo IQO09Pro	安卓14	✓	✓	✓	✓	×
红米K60 Ultra	安卓13	✓	✓	✓	×	✓
OPPO ace2	安卓12	✓	✓	✓	✓	×
小米6	安卓8	✓	✓	--	×	✓

• iOS

手机型号	系统	微信内置	Chrome	Safari
iphone13	ios16.3.1	✓	✓	✓
iphone14	ios17.5.1	✓	✓	✓
iphoneXS	ios15.3	✓	✓	✓
iphone11	ios17.5.1	✓	✓	✓
iphone15	ios17.5.1	✓	✓	✓
iphone 12	ios 16.3.1	✓	✓	✓
iphone12	ios14.6	✓	✓	✓

🔗 7. 常见问题

7.1 部分浏览器视频组件层级过高

移动端部分android浏览器对视频组件进行过特殊处理，导致视频组件悬浮于页面之上，该场景无法通过代码去解决

1. 如果只是层级过高，并不会直接全屏，那么可以使用半屏数字人或数字人头像大小，不在数字人界面上方叠加组件
2. 如果视频组件层级高且会直接全屏，那么可以直接拦截当前浏览器对数字人的使用

7.2 移动端集成，无法自动播放问题；（实现详见结尾demo）

移动端受限于浏览器策略，无法自动播放，需要一次用户交互行为（比如点击），常见处理方案

1. 进入数字人前有个引导页，页面上进行业务介绍与开始业务的按钮，点击按钮触发数字人的播放
2. 提前加载数字人iframe，然后将播放事件绑定在前置流程中的按钮上

相关配置：

entry=true

按钮触发方法：

dhlframe.playVideo(true);

注意：部分浏览器只是会继承上一次的用户交互时间，一段时间后也会失效，如果为了保障数字人的稳定性，建议前置加按钮触发

7.3 驱动数字人进行开场白话术，只有唇动无声音 （实现详见篇首demo）

受限于浏览器策略，无法自动有声音视频播放，sdk侧逻辑进行静音播放；

这时需要集成侧监听sdk抛出的rtcState消息localVideoMuted，引导用户完成交互操作后，

使用dhIframe.sendCommand({ subType: 'muteAudio', subContent: false });进行取消静音

7.4 部分iphone微信浏览器打开无数字人，部分安卓浏览器打开数字人不动

iphone微信浏览器无数字人原因：

微信限制，页面需要点击后才可播放视频(静音也不可自动播放)

iphone解决方案：

同6.4，集成侧监听sdk抛出的rtcState消息localVideoMuted/localVideoError，引导用户完成交互操作后，

使用dhIframe.sendCommand({ subType: 'muteAudio', subContent: false });进行取消静音

部分安卓浏览器数字人不动原因：

浏览器限制，页面中用户需要点击iframe区域，才可解除播放限制

解决方案：

诱导用户点击iframe区域后，调用dhIframe.sendCommand({ subType: 'muteAudio', subContent: false })

7.5 驱动数字人后，怎么中途打断数字人播报，以及如何监听数字人播报完成（实现详见篇首demo）

打断发送消息

```
dhIframe.sendMessage({action: 'TEXT_RENDER', body: '<interrupt></interrupt>'});
```

监听正常播放完成，首先指定TEXT_RENDER的requestId, 然后监听指定requestId的FINISHED; 监听播报因打断而提前完成，监听指定requestId的RENDER_INTERRUPTED;

7.6 数字人路数较少，如何节约资源，尽可能让资源在客户不用的时候尽快释放掉（实现详见篇首demo）

设置cp-inactiveDisconnectSec，cp-preAlertSec字段

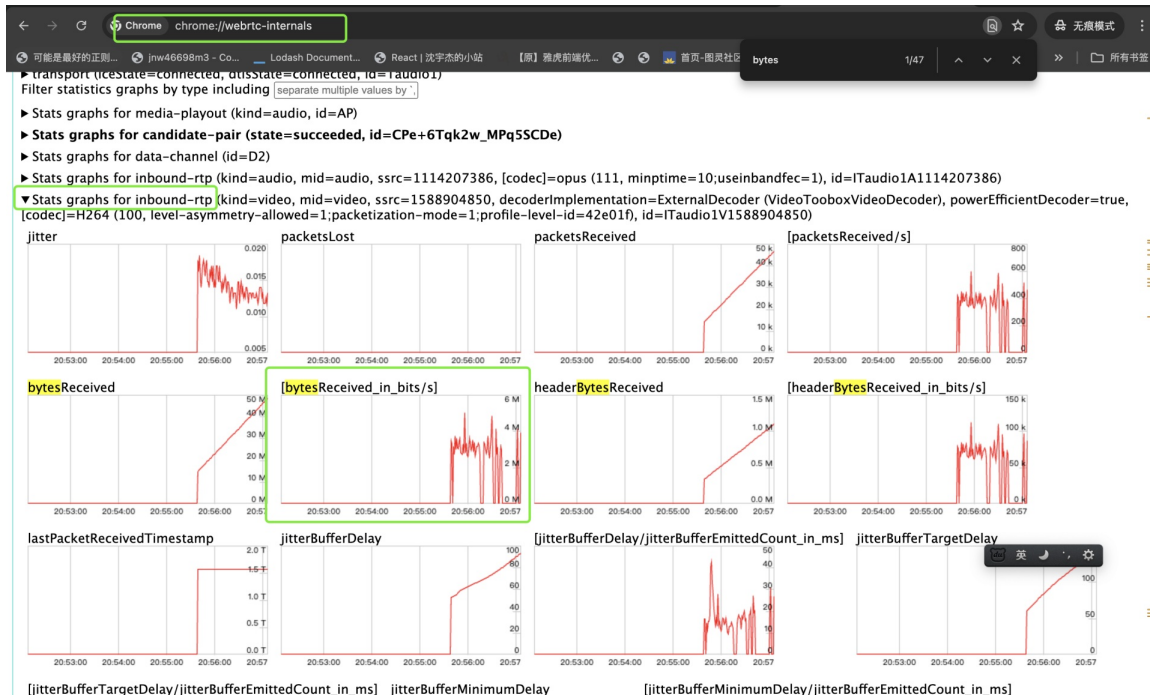
并监听DISCONNECT_ALERT进行适当的交互提示

监听TIMEOUT_EXIT消息进行销毁iframe操作

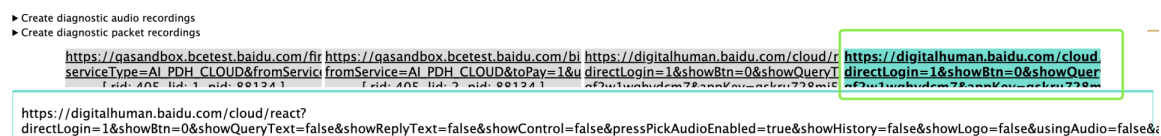
7.7 拉流没有画面

1. 请检查当前页面协议是否为HTTPS，若不是，请更换成HTTPS；

2. 请按照下图所示步骤继续排查客户端是否收到流



1. tab页开启h5-sdk保证已经拉流完成
2. 新开tab页输入chrome://webrtc-internals/
3. 点击列表中的最后一个链接



1. 页面全局搜索bytesReceived, 一直键入回车, 直到找到对应页面的[bytesReceived_in_bits/s]图表, 检查图标是否为空
2. 若该图表为空, 可能是浏览器插件导致, 为过滤浏览器插件影响, 请执行以下操作之一:
 1. 关闭所有浏览器插件, 重新打开页面
 2. 或打开chrome浏览器无痕窗口, 重新打开页面
 3. 或更换其他浏览器, 重新打开页面

目前已知会影响RTC正常拉流的浏览器插件: PureVPN Proxy;

如关闭所有浏览器插件并更换其他浏览器仍没有画面, 请提交工单与我们联系;

7.8 音频驱动效果差 接受的音频数据必须是 pcm无损音频编码格式, 需检测数据是否进行了压缩编码

如来源是音频 url, 检测 url 后缀是否是 pcm。如是 mp3, m4a, wav等格式则不支持

另外检查音频属性是否为支持"16k"、"16bit" 和 "单声道"

7.9 如何控制数字人大小 步骤 1: 通过设置 url 中参数mode=corp, resolutionWidth, resolutionHeight调整数字人原始视频流的分辨率, 以及设置合适的camerald

步骤 2: 为iframe添加外层div包裹, 并设置overflow: hidden

步骤 3: 通过iframe 的style中的width, height, top, left等属性, 调整数字人视频的大小到合适的尺寸, 位置处于预期的位置

```

<div className="iframe-wrapper">
<iframe
  id="digital-human-iframe"
  src="https://open.xiling.baidu.com/cloud/realtime?token=自行填写&initMode=noAudio&camerald=自行填写
&figureId=自行填写&resolutionWidth=1080&resolutionHeight=1920&mode=corp"
  allow="autoplay"
>
</div>
.iframe-wrapper {
  position: absolute;
  width: 100vw;
  height: 100vh;
  top: 0;
  left: 0;
  overflow: hidden;
}

#### digital-human-iframe {
  position: absolute;
  width: 140vw;
  height: 120vh;
  top: -20vh;
  left: -20vw;
}

```

7.10 demo运行报错如下

```

> realtime-digital-human-vue-demo@0.0.0 dev D:\jyc\新建文件夹 (3)\realtime-digital-human-vue-demo
> vite

(node:488832) UnhandledPromiseRejectionWarning: SyntaxError: Unexpected token "??="
    at Loader.moduleStrategy (internal/modules/esm/translators.js:145:18)
(Use "node --trace-warnings ..." to show where the warning was created)
(node:488832) UnhandledPromiseRejectionWarning: Unhandled promise rejection. This error originated either by throwing inside of an async function without a catch block, or by rejecting a promise which was
andled with .catch(). To terminate the node process on unhandled promise rejection, use the CLI flag "--unhandled-rejections-strict" (see https://nodejs.org/api/cli.html#cli_unhandled_rejections_mode). (r
on id: 1)
(node:488832) [DEP0018] DeprecationWarning: Unhandled promise rejections are deprecated. In the future, promise rejections that are not handled will terminate the Node.js process with a non-zero exit code

```

解决方案：使用yarn进行安装，并按照提示升级node版本

7.11 抠绿导致绿色配饰也被清除，自定义抠色参数 例如：人像无红色配饰，使用红背景扣红方案

步骤1. url删除autoChromaKey参数，为url添加扣色参数

```

chromaEffects =
{
  "version":2,"chromaKey":{"screen":[255,0,0],"similarity":90,"edgeShrink":1,"smoothness":210,"spill":100,"opacity":1,"cor
}

```

similarity: 默认70，所抠颜色与人像视频中颜色之间的相似度，值越高，删除的相似像素越多

edgeShrink: 默认为3，边缘收缩，边缘收缩的幅度系数，值越高，边缘黑边收缩幅度越大

screen: 为需要抠色的三原色数组，需要和步骤2设置的背景色一致

smoothness: 默认80，所抠颜色的平滑度，值越高，抠除边缘的平滑效果越强

步骤2. 设置颜色为[255,0,0] (#FF0000) 的纯色背景图。

url新增参数： backgroundImageUrl=<https://sdk-demo.bj.bcebos.com/red.jpg>

小程序版

1. 接入方法

1.1 确保小程序已拥有 live-player(实时播放音视频流)组件权限

申请并开通流媒体播放权限（符合开通【live-player | live-pusher】的资质文件）：

1.2 购买组件并申请数字人小程序插件权限

1. 在平台购买云渲染组件商品并绑定到应用，操作说明参考：[平台操作指引](#)

2. 申请数字人插件接入权限，插件申请地址：https://mp.weixin.qq.com/wxopen/pluginbasicprofile?action=intro&appid=wx702e7833d34cdcc5&token=&lang=zh_CN
3. 联系对接人审批数字人插件权限申请

1.3 引入插件代码包

使用者要在 app.json 中声明需要使用的插件

```
{
  "plugins": {
    "digitalhumanlib": {
      "version": "2.0.1",
      "provider": "wx702e7833d34cdcc5"
    }
  }
}
```

1.4 使用插件

和使用普通自定义组件的方式相仿，在 json 文件定义需要引入的自定义组件时，使用 plugin:// 协议指明插件的引用名和自定义组件名

```
{
  "usingComponents": {
    "digitalhuman-view": "plugin://digitalhumanlib/digitalHumanView"
  }
}
```

1.5 业务中使用

在 .wxml 中引入自定义组件

```
<digitalhuman-view
  id="digitalhuman"
  class="digitalhuman-view-container"
  settings="{{digitalHumanSetting}}"
  bind:onMessage="onMessage"
  bind:onStateChange="onStateChange"
/>
```

2. demo的使用

- 下载小程序demo, 下载地址：[**https://sdk-demo.bj.bcebos.com/demo_digitalHuman.zip](https://sdk-demo.bj.bcebos.com/demo_digitalHuman.zip)
- 使用微信开发者工具导入该文件，进行真机预览**

3. settings数字人配置项

3.1 基础配置

参数名	类型	必填	默认值	备注
token	string	是	-	平台购买交互组件商品，参考： 平台操作指南 ,生成鉴权凭证:生成方法查阅 接口通用说明 (token=Authorization)的鉴权参数生成部分
figureId	string	是	-	人像id，详细枚举参考： 3D数字人列表 、 2D精品数字人 、 2D小样本数字人
mode	string	是	-	用于针对特定应用场景快速批量设置关联配置的默认值，设置后关联配置无需填写，枚举类型： noAudio ：无拾音模式
resolutionWidth	number	否	720	数字人展示的宽度，必须是偶数，最小400，最大不超过 1080x1920、1920x1080，宽高比例需要和人像比例一致
resolutionHeight	number	否	1280	数字人展示的高度，必须是偶数，最小400，最大不超过 1080x1920、1920x1080，宽高比例需要和人像比例一致
backgroundImageUrl	string	否	-	设置数字人背景图片，推荐使用 BOS（百度对象存储服务）地址，其他地址可能会因为网络原因加载图片失败导致会话打开失败
x	number	否	0	水平位置，范围-0.5到0.5，取0居中，越大数字人越靠右
y	number	否	0	竖直位置，范围-1到1，取0为默认位置，越大数字人越靠上
z	number	否	1	缩放比例，范围0.5到1.5，取1原始比例，越小数字人越小。注意：y的取值是按数字人原始大小上下移动，当z取值大于1，既放大数字人时，y的值可以小于-1以实现更大范围下移。
cameraId	int	否	-	数字人预设机位 ID，控制数字人的位置和大小，建议使用和视频分辨率比例一致的机位：详细枚举参考： 3D数字人列表 、 2D精品数字人 、 2D小样本数字人

3.2 高级配置项

参数名	类型	默认值	备注
ttsPer	string	-	音色发音人ID，可用发音人列表参考： 公共音色库 ，或使用 克隆音色
ttsLan	string	auto	注：仅有LITE版克隆音色及公共音色库中的支持多语言的音色支持该参数。 • 支持参数列表：'Chinese', 'Chinese,Yue', 'English', 'Russian', 'Spanish', 'French', 'Portuguese', 'German', 'Turkish', 'Dutch', 'Ukrainian', 'Vietnamese', 'Indonesian', 'Japanese', 'Italian', 'Korean', 'Thai', 'Polish', 'Romanian', 'Greek', 'Czech', 'Finnish', 'Hindi', 'auto'
ttsPitch	number	5	音调，5是正常值，0-15的取值范围，越大声音越尖，默认值5（字面量需要是整数）
ttsSpeed	number	5	语速，5是正常值，0-15的取值范围。越大语速越快，默认值5（字面量需要是整数）
cp-autoAnimation	Boolean	false	是否开启自动添加数字人动作，只支持高精3D/2D人像
cp-inactiveDisconnectSec	int	180	长时间无交互（交互指发送 text render 让数字人播报）的断连提醒，单位秒，0 表示不开启。达到阈值时，服务端会发送 TIMEOUT_EXIT 提醒消息，页面在收到该消息时需销毁组件，避免数字人渲染资源浪费，目前数字人后端服务不会主动断开。
cp-preAlertSec	int	10	在开启长时间无交互断连提醒，且即将到达断连阈值时，后端会先发送一条 DISCONNECT_ALERT 预提醒消息，页面收到消息可以弹窗提醒用户。该参数用于设置断连前多少秒发送该提醒，比如无交互超时参数 cp-inactiveDisconnectSec 设置为 180（3 分钟），提醒参数 cp-preAlertSec 设置为 10（秒），则在用户最后一次交互后的 2 分 50 秒时会收到后端发送的 DISCONNECT_ALERT 提醒消息，在 3 分钟时收到 TIMEOUT_EXIT 断连消息。

4.1 文本驱动

代码样例详见handleSend

```
<digitalhuman-view
  id="digitalhuman"
  class="digitalhuman-view-container"
  settings="{{digitalHumanSetting}}"
  bind:onMessage="onMessage"
  bind:onStateChange="onStateChange"
/>

Page({
  data: {
    digitalHumanSetting: {...}
  },
  onLoad(options) {
    const digitalhuman = this.selectComponent('#digitalhuman');
  }

  handleSend(){
    // 文本驱动数字人
    this.digitalhuman.textRender({body: '这里是驱动数字人的文本', requestId: this.digitalhuman.getUuid()})
  }
});
```

requestId说明：可自定义构造requestId进行驱动数字人，驱动的回调消息会沿用自定义构造的requestId。

body说明：

TEXT_RENDER时body的可选值	备注
你好，我是数字人	播报内容 当 ttsLan 参数非'auto'时，请自行确认文本与所选语言是否一致，如不一致合成成功后会导致发音异常。
	打断当前数字人播报
<say-as type="telephone">110</say-as>	指定读法的高级播报内容，支持 在线SSML（语音合成标记语言）
开始说话停止5s后继续说话	支持播报中停顿
	音频播放
这段话不允许打断这段话不允许打断这段话不允许打断	播报不受打断的内容，后续textRender会在上句播报完成后开始播报，但仍可被interrupt标签打断

4.2 文本驱动回调

代码样例详见onMessage

```
<digitalhuman-view
  id="digitalhuman"
  class="digitalhuman-view-container"
  settings="{{digitalHumanSetting}}"
  bind:onMessage="onMessage"
  bind:onStateChange="onStateChange"
/>

Page({
  data: {
    digitalHumanSetting: {...}
  },
  onLoad(options) {
    const digitalhuman = this.selectComponent('#digitalhuman');
  }
  onMessage(message = {}) {
    const data = message.detail;
    const {action} = data;
    switch (action) {
      case 'RENDER_START':
        console.log('数字人驱动开始');
        break;
      case 'RENDER_COMPLETED':
        console.log('数字人驱动完成');
        break;
      default:
        break;
    }
  },
});
```

4.3 状态监听回调

代码样例详见onMessage

```
<digitalhuman-view
  id="digitalhuman"
  class="digitalhuman-view-container"
  settings="{{digitalHumanSetting}}"
  bind:onMessage="onMessage"
  bind:onStateChange="onStateChange"
/>

Page({
  data: {
    digitalHumanSetting: {...}
  },
  onLoad(options) {
    const digitalhuman = this.selectComponent('#digitalhuman');
  }
  onStateChange(e) {
    const {type, error = {}, loading} = e.detail;
    const {action, message, code} = error;
    if (type === 'INIT') {
      // 通过监听 INIT 消息，来实现数字人的开场白设置
      this.digitalhuman.textRender('数字人初始化完成，这是数字人的开场白播报，可以自行修改开场白内容');
    }
    if (type === 'ERROR'){
      console.log(type, error);
    }
  },
});
```

4.3.1 数字人初始化正常状态

type	说明
INIT	数字人第一次初始化完成
LOADING	数字人视频流是否加载中(type: 'LOADING', loading: true/false)

4.3.2 数字人初始化异常状态

type	code	错误信息	说明
ERROR	1001	服务器内部错误	
ERROR	1002	APP已失效	
ERROR	1003	APP token已过期	
ERROR	1004	目前线上用户较多，请您稍后再试	app 会话路数已经达到上限
ERROR	1005	授权信息不存在	figureId未授权不可用
ERROR	1009	App token非法	token填写错误
ERROR	1010	App key非法	
ERROR	1011	裁剪参数非法	仅2D数字人存在
ERROR	1012	参数不能为空	
ERROR	1014	分辨率/编码参数不合法	
ERROR	1015	无效的媒体信息	例如：传递的RTMP参数无效
ERROR	1202	正在连接或已连接	
ERROR	1204	解析请求失败	
ERROR	1206	读取json失败	
ERROR	2001	Fail to establish	建立渲染任务失败
ERROR	3001	目前线上用户较多，请您稍后再试	系统渲染资源不足导致
ERROR	23002	ttsPer非法	ttsPer音色不在公共音色库中

4.3.3 数字人驱动异常状态

type	code	错误信息	说明
ERROR	-1	body 非法	不支持 body 传& < >等特殊字符
ERROR	1013	body 非法	
ERROR	1208	body格式非法	body格式错误
ERROR	1201	操作尚未支持	

5. 数字人状态和指令控制

方法	备注
VERSION	获取当前sdk版本号
destroy	销毁数字人组件
pause(boolean)	true: 暂停近端视频流false: 播放并同步远端视频流
mediaMute(boolean)	true: 静音数字人false: 取消静音数字人

6. 常见问题

6.1 视频播放组件无权限，报错 operateLivePlayer:fail no permission

原生插件集成依赖：能够申请并开通流媒体播放权限（符合开通【live-player | live-pusher】的资质文件）：

需要在小程序开发管理上，申请并开启 live-player的权限



6.2 无法开通live-player权限如何在小程序中使用

可以使用web页面封装【云渲染组件】，再使用webview的方式进行嵌入小程序；

webview方式相较与原生小程序插件形式有如下特点

- 1. webview会自动铺满整个小程序页面。
- 2. 个人类型的小程序暂不支持使用
- 3. webview所支持api相较小程序插件有所不同，需要用户按需进行技术选型，相关参考微信技术文档：<https://developers.weixin.qq.com/miniprogram/dev/component/web-view.html>
- 4. webview兼容性同【云渲染组件】兼容性，存在无交互不可有声播放等问题，参见【云渲染组件】常见问题解决方案

webview接入方式，添加校验文件具体步骤如下

- 1. 登录微信公众平台，进入小程序管理页面。
- 2. 在“开发设置”中找到“业务域名”部分，点击“修改”按钮。
- 3. 在“合法域名”列表中，点击“添加”按钮，输入iframe域名：<https://open.xiling.baidu.com>和业务域名
- 4. 选择文件验证或DNS验证，按照提示完成验证。



- 1. 一旦验证成功，<https://open.xiling.baidu.com> 和业务域名就会出现在“合法域名”列表中，你的小程序就可以通过这个域名访问服务器了。

2. 联系对接人员，将校验文件发给对接人员，待检验文件成功放到数字人服务的根目录上后，webview接入准备工作完成

交互会话管理接口

接口说明

接口调用域名：<https://open.xiling.baidu.com>

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

本文档提供的接口可以方便用户对交互会话进行管理，在用户发现云渲染交互资源并发不足无法拉流时，可以通过对云渲染交互会话进行查询，获取会话的相关信息，以帮助区分不同的会话，并可以主动断连部分会话以释放资源。

会话管理接口支持：

- 1.获取当前app id下进行中的云渲染交互会话列表
- 2.主动关闭交互会话

接口列表

查询会话列表

接口说明

查询当前 app id 下进行中的云渲染交互会话。

接口协议

GET /api/digitalhuman/open/v1/session/interaction/list

返回参数

参数	类型	说明
sessions	list	会话列表
-- sessionId	string	会话ID
-- brtcPreviewUrl	string	会话视频流预览地址，用于区分会话
-- startTime	string	会话开始时间

返回示例

```
{
  "code": 0,
  "message": {
    "global": "success"
  },
  "result": {
    "sessions": [
      {
        "sessionId": "r-xxx",
        "brtcPreviewUrl": "https://brtc.cdn.bcebos.com/brtc.html?appid=xxx&roomname=xxxxx&listener_only=1",
        "startTime": "2025-04-14T14:10:55+08:00"
      }
    ]
  },
  "requestId": "5b0cdc66-c6a8-4d85-a92e-a80b467a2fa4",
  "success": true
}
```

错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
14001	内部服务异常，请稍后再试
14002	内部服务异常，网络拥堵

🔗 主动关闭会话

接口说明

主动关闭交互会话。

接口协议

DELETE /api/digitalhuman/open/v1/session/interaction?sessionId={sessionId}

请求参数

参数	类型	必填	说明
sessionId	list	是	需要关闭会话的会话ID

请求示例

```
DELETE /api/digitalhuman/open/v1/session/interaction?sessionId=r-xxx
```

返回示例

```
{
  "code": 0,
  "message": {
    "global": "OK"
  },
  "result": null,
  "requestId": null,
  "success": true
}
```

错误码

错误码	描述
0	正常返回
4911	找不到app信息，请确认appId是否输入正确
4913	无法访问API，可能是app没有绑定对应的组件，或url输入错误，或访问的人像不可用
10001	签名校验失败
10002	签名信息为空
10003	签名格式错误
10004	未识别错误的通用错误码
10005	请求体JSON解析失败，请确认是否是合法的JSON格式
10006	参数校验不通过
10011	没有购买对应的商品
14001	内部服务异常，请稍后再试
14002	内部服务异常，网络拥堵

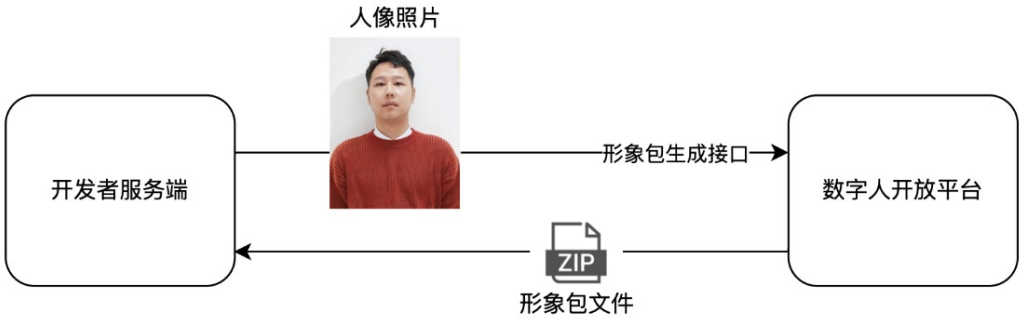
端渲染交互

照片数字人交互SDK-H5版

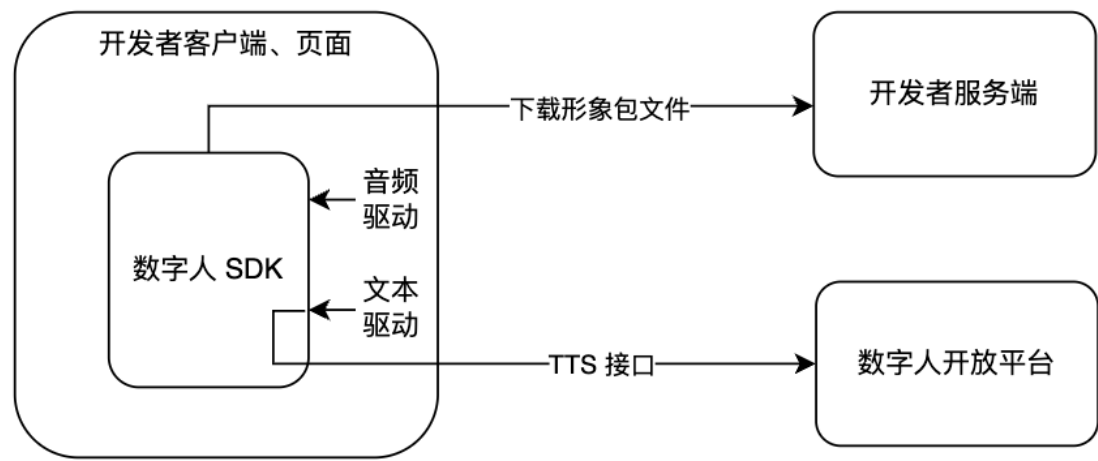
1. SDK运行原理

1.1. 使用流程

1、形象定制：服务端调用数字人开放平台“形象定制-照片数字人”接口，上传人像照片获取形象包文件。形象定制参考[照片数字人形象定制接口文档](#)，已生成形象包需要通过SDK驱动使用，不可单独使用。



2、SDK调用：本文档为端渲染交互组件-照片数字人SDK文档，客户端或页面集成集成SDK加载形象包，本地驱动渲染数字人。支持音频和文本驱动。音频可以使用真人音频或自行接入的第三方TTS的音频。文本驱动SDK 内部会调用数字人开放平台TTS 服务驱动数字人。



1.2 购买入口

[组件商店](#)

2. 接入方法

2.1 安装依赖包

```
npm i @bddh/starling-dh-picture-client
```

```
或 yarn add @bddh/starling-dh-picture-client
```

SDK 包不定期进行升级修复边缘 case（截止至 2025.1.14，最新版本为1.0.44），如遇问题，请先更新sdk 包版本

2.2 代码中使用

```
import PictureHuman from '@bddh/starling-dh-picture-client';

const instance = new PictureHuman({
  modelUrl: '形象 zip 包',
  licenseUrl: 'license 的 url',
  licenseKey: 'license 的 key',
  statusCallback: (status: object) => console.info(JSON.stringify(status)),
  canvasId: '#canvas-pic' // 跟下文的 id 一致，添加#
});

return (
  <canvas id="canvas-pic"></canvas>
)
```

3.SDK 兼容性列表

支持主流浏览器，详细兼容见下表

端\浏览器	Chrome	Edge	Safari	Opera	UC	百度（手百）	Firefox	IE	QQ
移动端-安卓	123+	——	——	——	15.5+	13.52+	——	——	14.9+
移动端-iOS	126	——	13.3.1+	——	16.6.0	13.60	——	——	15.2
PC端-Mac	108+	——	14.1.1+	111	——	——	116+	不兼容	5.0.4
PC端-Windows	108+	126	——	111	——	——	115+	不兼容	12.6

支持机型信息：

- 安卓：最低支持安卓10 CPU骁龙730G/骁龙855/联发科天玑900/天玑820(同等性能的芯片)及以上

- iOS：最低支持iphonexs 系统版本13.3以及以上。

包大小和相关时延

SDK大小	12.4M + CDN引擎资源包4.3MB
自定义人像包大小	6MB左右
音频驱动时延	100ms 左右
文本驱动时延	0.3s~1.5s(跟第一个标点前文字数量相关)

4.demo 使用

4.1 demo说明

此demo分别使用react，vue两种框架。

针对如何发送消息驱动数字人, 发送流式音频驱动数字人，以及数字人的打断，状态的检测，给予了相应的实现

其中文本驱动的 token需要自行生产生成

4.2 demo下载

React版本，下载地址<https://sdk-demo.bj.bcebos.com/picture-human-client-react-demo.zip>

Vue版本，下载地址<https://sdk-demo.bj.bcebos.com/picture-human-client-vue-demo.zip>

4.3 demo运行

1. 命令行执行yarn，安装所需依赖包 (根据提示使用对应的 node 版本)
2. 命令行执行npm run dev，启动项目
3. 将控制台输出的ip地址，复制到浏览器地址栏进行项目预览（目前cors只可使用 localhost:8080进行预览）

5. 初始化配置

参数	必填	类型	含义
modelUrl	是	string	模型包地址
licenseUrl	是	string	模型鉴权文件地址
licenseKey	是	string	模型鉴权 key
statusCallback	否	function	返回 5.1的状态信息

5.1 状态消息

type	value	含义
status	{ code: 0, value: 'PENDING', progress: number}	模型加载中, progress为加载进度，值为 0-100
status	{ code: 0 value: 'SUCCESS', progress: 100}	模型加载完成
code	--	鉴权相关状态码, 枚举见 5.2 表

5.2 初始化错误码

code	含义
1	license未初始化
2	license数据解密失败
3	license数据格式错误
4	license-key(api-key)校验错误
5	算法ID校验错误
6	MD5校验错误
7	设备ID校验错误
8	域名校验错误
9	过期时间不正确
10	功能未授权
11	授权已过期
12	本地文件读取失败
13	远程数据拉取失败
14	本地时间校验错误
1001	构造函数的ready属性不是函数
1002	构造函数的config属性不是对象
1003	构造函数的licenseKey属性不是字符串
1004	构造函数的license属性不是字符串
1005	多级嵌套iframe不被支持
1101	初始化webgl的上下文失败
1201	replaceVisualizeBody方法的参数不合法
1301	init方法的参数不合法
1401	构造函数的config属性不合法
1501	创建虚拟路径失败
1502	创建虚拟文件失败
1503	创建鉴权虚拟文件失败
1601	初始化模型失败
2001	无用户交互，无法自动播放音频

6.驱动与回调

6.1 音频驱动

```
instance.audioRender({
  audioBuffer: 'arrayBuffer 格式的数据',
  first: true,
  last: true,
  sample: 16000,
  interrupted: true
  finishListener: err => console.info('播报完成', err)
});
```

参数	必填	类型	默认值	含义
audioBuffer	是	ArrayBuffer	-	播报音频（必须是 PCM 无损音频编码格式，MP3属于有损压缩，其数据不可驱动**）
sample	是	16000	-	音频数据的采样率，目前只支持16K
first	否	boolean	true	流式音频驱动:第一条音频数据传first= true 完整音频驱动：first=true或不传
last	否	boolean	true	流式音频驱动:最后一条音频数据传last= true 完整音频驱动：last=true或不传
interrupted	否	boolean	false	音频驱动是否打断正在播报的内容 false: 不打断当前播报内容 true: 停止当前播报内容，直接播报驱动音频
finishListener	否	(err?: Error) => void	-	驱动结束回调，报错失败会通过该回调返回Error 参见6.4驱动错误 Error

🔗 6.2 文本驱动

```
instance.textRender({
  msg: {
    token: '填写 appKey, appId 生成的鉴权 token',
    text: '使用文本驱动数字人的播报文本',
    tts: {
      per: 5116,
      spd: 5,
      pit: 5,
      vol: 5
    },
    interrupted: true
  },
  finishListener: err => console.info('播报完成', err)
});
```

参数	必填	类型	含义
text	是	string	播报文本, 支持 在线 SSML(语音合成标记语言) ; 不可包含& < >特殊字符. •当使用LITE版克隆音色及公共音色库中支持多语言的音色时, 支持自动识别输入语言来进行播报。例：当输入俄文时, 使用俄语播报。
token	是	string	鉴权 token 鉴权参数生成
finishListener	否	(err?: Error) => void	驱动结束回调, 报错失败会通过该回调返回Error 参见6.4驱动错误 Error
tts	是	{ per: number string; spd: number; pit: number; vol: number }	per 发音人, 取值参考 公共音色库 ; spd 语速 5是正常值, 0-15的取值范围。越大语速越快, 默认值5 (字面量需要是整数) pit 语调 5是正常值, 0-15的取值范围, 越大声音越尖, 默认值5 (字面量需要是整数) vol 音量 5是正常值, 0-15的取值范围, 越大音量越大, 默认值5 (字面量需要是整数)
interrupted	否 (默认 false)	boolean	音频驱动是否打断正在播报的内容 false: 不打断当前播报内容, 多用于流式文本驱动 true: 停止当前播报内容, 直接播报驱动音频, 多用于整段文本驱动
last	否 (默认 true)	boolean	流式文本驱动: 最后一条文本数据传last= true 其余文本数据传last=false 整段文本驱动: last=true或不传

🔗 6.3 指令驱动

方法	返回值	使用方法	含义
isPlay	boolean	humanInstance.isPlaying()	判断数字人是否正在播报
play	promise	await humanInstance.play()	开始播报
isPause	boolean	humanInstance.isPause()	判断数字人是否正在暂停
pause	promise	await humanInstance.pause()	播报暂停
isStop	boolean	humanInstance.isStop()	判断数字人已经播报完成
stop	promise	await humanInstance.stop()	终止播报
finish	void	humanInstance.finish	销毁数字人

🔗 6.4 驱动错误 Error: {code, message}

code	含义
1702	音频驱动数据，audioBuffer异常
1701	sample，采样率设置非法
2001	无用户交互，无法自动播放音频
TTS_1003	鉴权信息已过期
TTS_1004	超出并发上限，不可使用 tts
TTS_1005	鉴权信息不存在
TTS_1007	过期时间格式非法
TTS_1008	鉴权信息非法
TTS_1009	App key非法
TTS_1010	App token非法
TTS_4001	文本驱动时，模型加载未完成

7. 常见问题

7.1 浏览器置于后台播报停止

因为使用WebAssembly技术，程序运行会在浏览器置于后台或者标签页失活后自动停止，属于正常表现；

如需切换回页面，进行重播，可以设置时间监听，重新发送文本或音频驱动指令

7.2 js 和 license跨域cors报错

获取到的 zip 包和 license有两种处理方式来解决 cors报错；

1. 上传到 cdn，为 cdn 设置 cors白名单
2. 保证zip, license资源域名同 sdk集成的业务同域名。

7.3 音频使用 ArrayBuffer数据进行驱动，但是效果不好

sdk 照片数字人接受的音频数据必须是 pcm无损音频编码格式，需检测数据是否进行了压缩编码

如来源是音频 url，检测 url 后缀是否是 pcm。如是 mp3，m4a, wav等格式则不支持

7.4 文本数字人播报一半后卡住不动

1. 流式文本，检查消息内容是否为""空字符串
2. 流式文本，驱动检查是否最后有last: true标识

7.5 数字人播报只有口型没有声音

根因：浏览器在无用户交互（交互：用户点击，拖拽等）的情况下，播放有声音视频，会自动开启静音；

解决方案：在数字人播报之前，引导用户进行点击等行为操作。无法有声播放得驱动回调：{code: 2001 msg: 无用户交互，无法自动播放音频}

2D数字人交互SDK-Android版

一、产品介绍

提供端渲染交互SDK-2D小样本数字人，可以通过文本或PCM音频数据完成对虚拟人实时驱动。

1、适用场景特点 部署成本低：支持低成本快速集成，人像渲染及唇形推理无需依赖云端服务；

网络依赖小：可落地在金融、政务等多种场景，适用线下硬件设备。

2、核心功能 公共形象租赁：提供半身、全身的2D小样本数字人形象，覆盖多种场景，支持客户快速集成使用；
专属形象定制：支持定制专属的2D小样本数字人形象；
实时播报渲染：支持数字人实时播报，可应用在多种线下终端设备对话问答场景；
性能指标：在3588及以上性能设备。

3、SDK下载链接 SDK下载地址：[https://digital-human-](https://digital-human-cloud.bj.bcebos.com/open/sdk/2d/baidu_digital_SDK_V1.5.7.25012214_Release_20250702_1435.zip)

[cloud.bj.bcebos.com/open/sdk/2d/baidu_digital_SDK_V1.5.7.25012214_Release_20250702_1435.zip](https://digital-human-cloud.bj.bcebos.com/open/sdk/2d/baidu_digital_SDK_V1.5.7.25012214_Release_20250702_1435.zip)

4、SDK的详细使用 **4.1、在初始化页面：**第一次进入demo之后，请填写appKey、secretKey和时长包，如需文本驱动请填写ttsId和ttsKey，然后先点击确认使用按钮（注意：时长包的确认使用，如果已经续期成功就不用点击了），再点击初始化SDK按钮，最后点击SDK授权+启动数字人（SURFACEVIEW），其他按钮请不要点击了。

4.2、进入数字人页面：有两种播报模式，一种是普通播报，一种是流式播报，可点击右上角进行切换（流式播报切换到普通播报之前，需要先关闭流式播报）。

4.2.1、普通播报（1）播报功能的使用：需要点击播放长文本或播报短文本或播放PCM，其中文本驱动是播放长文本和播报短文本，播放PCM播放的是本地PCM文件。

普通播报一次只可播报一条数据（在语音播报期间不可以点击播放长文本或播报短文本或播放PCM），可以插入播报，插入播报有插入PCM播报和插入TXT播报（**注意：如未播报语音，插入pcm播报会失败的**），插入的播报会立即在当前播报中插入播报。

（2）切换形象：如需切换其他形象必须要有一个定制的形象包《具体可参考本文档8.1、定制人像包替换本地数字人像包》（注意：在播报中不可切换形象）。

（3）暂停/继续播报：暂停播报可使正在播报的播报暂停，继续播报可使已经暂停的播报恢复播报。

（4）打断播报：打断播报可使所有的普通播报（包括插入的所有播报）停止，且不可恢复。

（5）能力释放：可释放资源并且退出数字人页面。

注意：语音播报期间不可以点击播放长文本或播报短文本或播放PCM。

4.2.2、流式播报 流式播报就是可持续的播报，可把新的播报追加到当前播报后边

（1）播报功能的使用：首先点击打开流式播报按钮打开流式播报，流式播报需要点击持续PCM或持续TXT，其中文本驱动是持续TXT，持续PCM播放的是本地PCM文件。

流式播报就是可持续的播报，可以持续播报多条数据，就是把新的播报追加到当前播报后边（在播报期间点击持续PCM或持续TXT就是把新的PCM播报或TXT播报追加在当前播报的后边进行播报了）。

在语音播报期间可以插入播报，插入播报的时候需要点击开始插播按钮，这时开始插播按钮变成停止插播按钮了（**插播期间不支持点击持续PCM和持续TXT按钮，请勿在插播期间进行持续播报**），当前播报就暂停了，插入播报有插入PCM播报和插入TXT播报，插入的播报会立即播报出来，当不需要插播的时候，点击停止插播按钮，这时可以点击持续PCM和持续TXT按钮追加播报了。

（2）切换形象：如需切换其他形象必须要有一个定制的形象包《具体可参考本文档8.1、定制人像包替换本地数字人像包》（注意：在播报中不可切换形象）。

（3）暂停/继续播报：暂停播报可使正在播报的播报暂停，继续播报可使已经暂停的播报恢复播报。

（4）打断播报：打断播报可使所有的流式播报（包括插入的所有播报）停止，不可恢复。

（5）能力释放：可释放资源并且退出数字人页面。

注意：插播期间不支持点击持续PCM和持续TXT按钮，请勿在插播期间进行持续播报。

二、SDK集成

1、支持的系统和硬件版本

请注意：这是对于开发人员来说的，请开发人员按照此要求进行集成

建议rk3588及以上性能安卓板卡

系统	支持Android 7.0+（API Level 24）及以上
CPU架构	armeabi-v7a, arm64-v8a
硬件要求	性能要求：硬件设备CPU 8核及以上，内存8G及以上，可用存储空间500MB及以上。
网络	支持WIF及移动网络。如果使用云端TTS（以文本方式驱动数字人），设备带宽（用于数字人的实际带宽）期望10mbps及以上。
开发IDE	Android Studio Flamingo 2022.2.1 Patch 1
内存要求	SDK可用于数字人的内存 >= 400MB

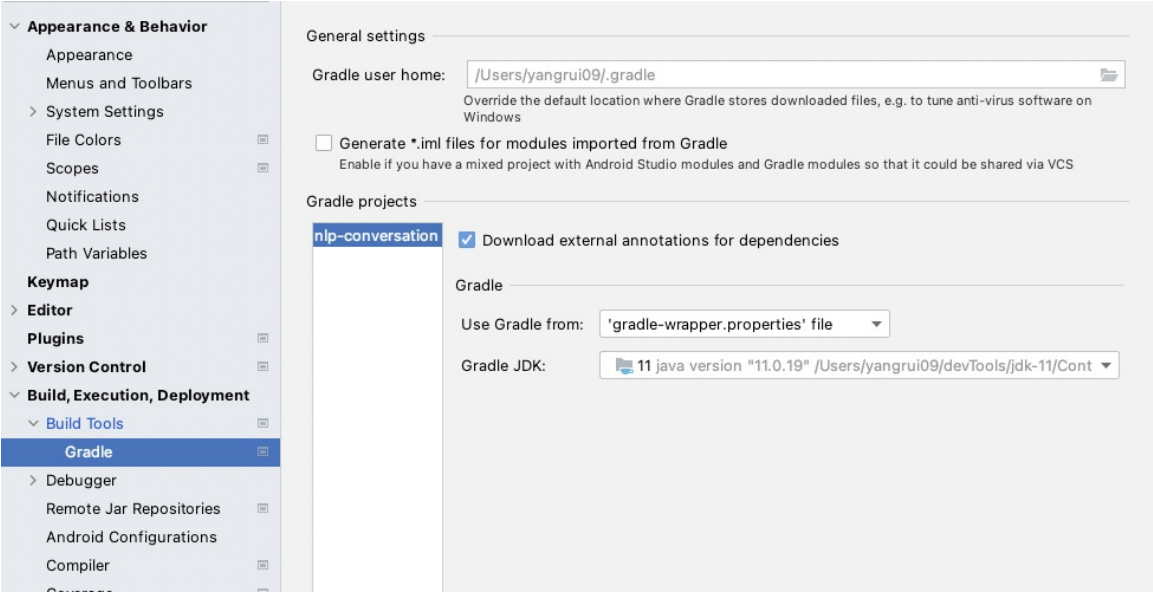
2、SDK集成 引入sdk aar 包：module_core-jnidev-release.aar，module_virtual-release.aar

2.1、app目录新建libs目录，放入aar包，在build.gradle中增加配置如下

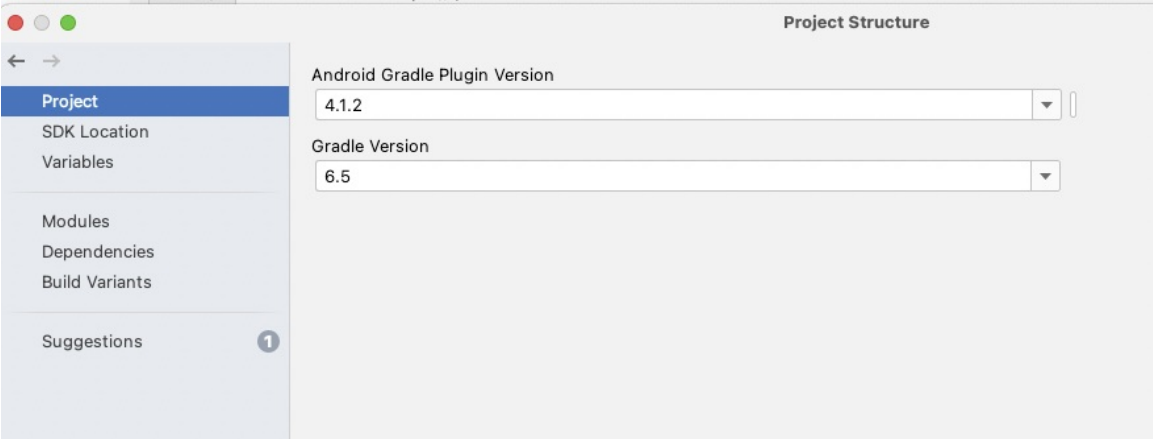
```
implementation fileTree(dir: "libs", include: ["*.jar"])
implementation files('libs/module_core-jnidev-release.aar')
implementation files('libs/module_virtual-release.aar')

implementation 'com.google.code.gson:gson:2.8.9'
implementation("com.squareup.okhttp3:okhttp:4.11.0")
```

2.2、android studio 配置 2.2.1、设置-> Build, Execution, Deployment-> Build Tools -> Grale -> Gradle JDK, 设置java 11



2.2.2、工程设置（mac 下，快捷键command+;) Android Gradle plugin Verison



2.2.3、Android SDK & java 版本设置

以下配置为建议配置，其他版本可能存在问题。



3、系统权限配置以及申请 3.1、权限配置

```
<uses-permission android:name="android.permission.MODIFY_AUDIO_SETTINGS" />
<uses-permission android:name="android.permission.CHANGE_WIFI_STATE" />
<uses-permission android:name="android.permission.ACCESS_NETWORK_STATE" />
<uses-permission android:name="android.permission.ACCESS_WIFI_STATE" />
<uses-permission android:name="android.permission.INTERNET" />
<uses-permission android:name="android.permission.READ_PHONE_STATE" />
<uses-permission android:name="android.permission.WRITE_EXTERNAL_STORAGE" />
<uses-permission android:name="android.permission.READ_EXTERNAL_STORAGE" />
<uses-permission android:name="android.permission.MANAGE_EXTERNAL_STORAGE" />
```

3.2、权限申请

请参考demo BaseActivity示例，所列举的权限为最低权限要求，如果缺少某一个，会导致sdk出现异常。

三、SDK调用及API说明

1、初始化相关 1.1、初始化SDK- 设备授权

初始化是指数字人资源以及形象资源初始化， 调用方式参考接口相关文档

appKey、seceretKey为一台设备独有，一个设备要使用数字人必须要有一对appKey、seceretKey，使用这两个参数就可以使用音频驱动了。

如需文本驱动，就必须要有对ttsId、ttsKey，ttsId和ttsKey为组件平台创建的tts合成应用的appld和appKey，
tts音色参考：https://cloud.baidu.com/doc/AI_DH/s/plyy6xhi0。

```
int init(Context context, VirtualInitParam virtualInitParam);
```

参数说明：

参数名	取值	说明
context	Context	非空
virtualInitParam	VirtualInitParam	非空

VirtualInitParam的4个参数说明

参数名	取值	说明
appKey	appKey	非空，这两个值，由百度提供，为一台设备的独有。请向百度申请。
seceretKey	seceretKey	
ttsId	ttsId	如需文本驱动，就必须要有ttsId、ttsKey。请在组件平台创建tts合成应用，
ttsKey	ttsKey	获取方式请参考文档： https://cloud.baidu.com/doc/AI_DH/s/plyy6xhi0 。

返回参数说明：

code	说明
0	初始化成功
-1000	Context为null
-1001	AppKey为空
-1002	SecretKey为空
-1003	AppKey长度错误
-1004	SecretKey长度错误
-1005	设置文件路径失败
-1006	引擎状态错误，重复初始化
-1007	callback为null
-1008	项目授权请求结果错误
-1009	ProjectAK为空
-1010	ProjectSK为空
-1011	ProjectAK长度错误
-1012	ProjectSK长度错误
-1013	项目授权时Context为null
-1014	项目授权时VirtualAuth为null
-1015	VirtualInitParam为空
-1016	ttsId为空
-1017	ttsKey为空

调用示例

```
public class MainApplication extends Application {
    @Override
    protected void attachBaseContext(Context base) {
        super.attachBaseContext(base);
        VirtualFactory.getEngine().init(this, Config.APP_KEY, Config.SECERET_KEY);
    }
}
```

1.2、获取自定义模型地址

用于获取非内置模型文件地址， 1、通过该接口获取地址，2、将所需模型推送至该地址目录下

```
String getThirdVirtualModelPath()
```

调用示例

```
String modelPath = VirtualFactory.getEngine().getThirdVirtualModelPath()
```

1.3、更新ttsId和ttsKey参数

ttsId和ttsKey为组件平台创建的tts合成应用的appId和appKey，获取方式请参考文档：https://cloud.baidu.com/doc/AI_DH/s/plyy6xhi0。

```
boolean updateTtsParam(String ttsId, String ttsKey);
```

参数说明：

参数名	取值	说明
ttsId	ttsId	如需文本驱动，就必须要有ttsId、ttsKey。请在组件平台创建tts合成应用，获取方式请参考文档： https://cloud.baidu.com/doc/AI_DH/s/plyy6xhi0 。
ttsKey	ttsKey	

调用示例

```
VirtualFactory.getEngine().updateTtsParam(ttsId, ttsKey);
```

2、SDK授权

授权是针对数字人功能能力授权，如果不授权或者授权不通过将无法使用数字人功能，授权方式可参考文档

2.1、获取SDK授权状态

```
public synchronized boolean getOAuthState(final Context context, final AuthStateCallBack authStateCallback)
```

参数说明

参数名	类型	取值	说明
context	Context		上下文
authStateCallback	AuthStateCallBack	code：0：成功，其他：失败 subCode: 后端返回状态码 activeTime：生效时间，异常时返回0 expirationTime：到期时间，异常时返回0 msg：提示消息	subCode为后端返回状态码。由于该码可能会有变动，所以不可作为业务逻辑判断使用
返回值	boolean	true：获取成功 false：获取失败	

code取值：

code	说明
0	续期成功
-1001	获取状态失败，SDK状态异常
-1002	获取状态失败，后台返回为空
-1003	获取状态失败，后台状态不成功
-1004	获取状态失败，没有返回license
-1005	获取状态失败，一般是网络异常

调用示例

```
VirtualAuth vhAuth = new VirtualAuth()
vhAuth.getOauthState(mContext, (code, subCode, activeTime, expirationTime, msg) -> {
    // 相关业务处理逻辑
});
```

2.2、使用时长码

当使用新的 **appKey**，**seceretKey** 时必须进行激活，激活方式为使用时长码激活，时长码会随着appKey，seceretKey 提供，使用项目授权方式时不需要时长码激活。

```
public synchronized boolean renewExpirationDate(final Context context, String activateCode,
                                                final AuthRenewCallBack renewCallBack)
```

参数定义：

参数名	类型	取值	说明
context	Context	非空	
activateCode	String	非空	有效的激活码
renewCallBack	AuthRenewCallBack	code：0：成功，其他：失败 subCode：后端返回状态码 expirationTime：有效的时候取有效值，异常时返回0 msg：提示消息	subCode为后端返回状态码。由于该码可能会有变动，所以不可作为业务逻辑判断使用
返回值	boolean	true：获取成功 false：获取失败	

code取值如下：

取值	说明
0	续期成功
-1000	续期失败，发生异常
-1001	续期失败，SDK状态异常
-1002	续期失败，后台返回为空
-1003	续期失败，后台状态不成功

调用示例

```
VirtualAuth vhAuth = new VirtualAuth()
vhAuth.renewExpirationDate(mContext, inputCode, (code, subCode, expirationTime, msg) -> {

});
```

2.3、SDK授权初始化

数字人能力授权， 成功后可执行数字人相关逻辑

```
public synchronized boolean initLicenseBatch(final Context context, final AuthCallBack authCallback)
```

参数说明

参数名	类型	取值	说明
context	Context		上下文
authCallba ck	AuthCallBa ck	code：0：成功，其他： 失败 subCode：后端返回状态 码 msg：提示消息	subCode为后端返回状态码。由于该码可能会有变动，所以不可作为业务逻辑判断使用

code取值(目前取值如下，未来会增加)：

取值	说明
0	授权成功
-1000	授权失败
-1001	tts授权失败
-1002	设备ID失败
-1003	引擎状态异常
-1004	模型文件异常

调用示例

```
VirtualAuth vhAuth = new VirtualAuth()
vhAuth.initLicenseBatch(mContext, (code, subCode, msg) -> {

});
```

2.4、换绑授权码设备

当一个授权码还存在有效期，但是设备坏了的情况下，可以使用这个接口进行授权设备的换绑，使用项目授权方式时不可进行换绑。

换绑接口有次数限制，非特殊情况不宜调用

```
public synchronized boolean changeDeviceBind(final Context context, final AuthStateCallBack callBack)
```

参数说明

参数名	类型	取值	说明
context	Context		上下文
callBack	AuthStateCallback	code：0：成功，其他：失败 subCode：后端返回状态码 msg：提示消息	subCode为后端返回状态码。由于该码可能会有变动，所以不可作为业务逻辑判断使用

code取值

取值	说明
0	授权成功
-1001	切换失败，SDK状态异常
-1002	切换失败，后台返回为空
-1003	切换失败，后台状态不成功
-1004	切换失败，一般是网络异常(也可能是调用方代码异常了)

调用示例

```
VirtualAuth vhAuth = new VirtualAuth()  
vhAuth.changeDeviceBind(mContext, (code, subCode, msg) -> {  
  
});
```

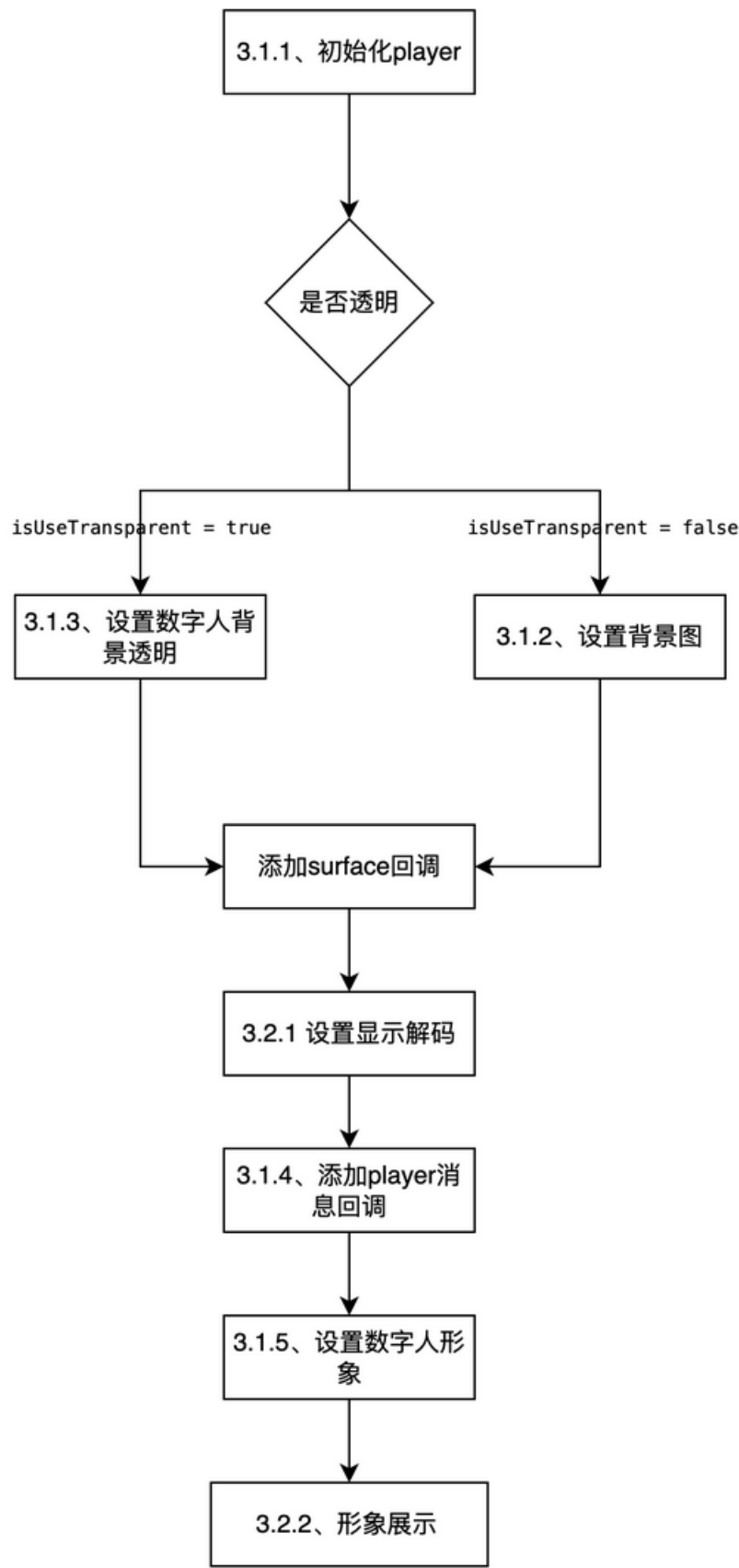
3、数字人形象

数字人是通过Player 对象进行管理，每次初始化要求为一个Player对象，用于调用数字人相关操作接口，具体接口参考接口文档，数字人显示部分可参考demo MainActivity.java 示例代码

静默态说明：当开启静默态时（isCloseSilent = false）未进行播报时可以展示形象和背景；
当未开启静默态时（isCloseSilent = true）未播报时不会展示形象和背景，播报时会展示形象和背景，播报完成后会定在最后一帧，如果这时退出后台再次进入则会进入默认态（不会展示形象和背景）

3.1、数字人形象初始化 surface 模式初始化流程

surface模式下初始化调用接口流程



3.1.1、初始化Player

请注意：初始化Player后，请先设置数字人形象《具体可参考本文档3.1.5、设置数字人形象》。

```
public static Player getPlayer(Context context)
```


参数名	类型	取值	说明
context	Context		上下文
返回值	Player		用于Player 相关接口后续调用

调用示例

```
Player player = VirtualFactory.getPlayer(mContext);
```

3.1.2、设置背景图

如果需要背景为透明状态则无需调用此接口设置背景，直接参考 3.1.3 设置即可

```
boolean setBackground(Bitmap bitmap)
```

参数名	类型	取值	说明
bitmap	Bitmap		背景图片

调用示例

```
player.setBackground(BitmapFactory.decodeResource(getResources(), R.mipmap.back));
```

3.1.3、设置数字人背景透明

直接设置SurfaceView 为透明状态即可

```
mSurfaceView.getHolder().setFormat(PixelFormat.TRANSLUCENT);
mSurfaceView.setZOrderOnTop(true);
```

3.1.4、添加player 消息回调

```
boolean addCallback(Callback callback);
```

参数名	类型	取值	说明
callback	Callback		设置回调

Callback 回调说明

参数名	取值	说明
void onVirtualHumanStart()		数字人播放开始
void onVirtualHumanFinish()		数字人播放结束
void onVirtualHumanData(VirtualHumanDataBean frameBean)	frameBean.getAudioStreams() PCM 音频数据 frameBean.getVideoStreams() 数字人形象RGBA 数据	数字人音频数据和形象RGBA 数据, 使用前需做判空处理
void onRemainBuffer(int remainBytes, intremainInsertBytes)	remainBytes : 剩余未播放的buffer remainInsertBytes : 剩余未播放插播的buffer数据	数字人剩余未播放的buffer长度
void onVirtualHumanPause()		暂停播放
void onPcmVolumeError(int code, String message, float vol)	code: -1 : 扩大失败 message: 错误信息 vol : 1倍-10倍 : 当前设置的倍数	幅值放大后出现截幅失真会进行触发
void onError(int code, intsubCode, String msg)	code : -1000 : tts数据异常 -1001: 播放器状态不合适 subCode : 详情参考 7.3错误信息参数说明 msg: 错误信息	出错消息

调用示例

```
player.addCallback(new Player.Callback() {  
    @Override  
    public void onVirtualHumanStart() {  
  
    }  
  
    @Override  
    public void onVirtualHumanFinish() {  
  
    }  
  
    @Override  
    public void onVirtualHumanData(VirtualHumanDataBean frameBean) {  
  
    }  
  
    @Override  
    public void onRemainBuffer(int remainBytes, int remainInsertBytes) {  
  
    }  
  
    @Override  
    public void onPcmVolumeError(int code, String message, float vol) {  
  
    }  
  
    @Override  
    public void onVirtualHumanPause() {  
  
    }  
  
    @Override  
    public void onError(int code, int subCode, String msg) {  
  
    }  
});
```

3.1.5、设置数字人形象

```
void setVirtualModel(Model model, DisplayMode displayMode, ChangeVirtualModelCallback callback)
```

参数名	类型	取值	说明
model	Model	visualize：形象ID，visualize值默认是10001，内置的默认人像是"慕晴"，如加载定制人像包，visualize值则要大于0且不等于20，也不能等于10001-11000中的数，还不能重复，visualize值为数字人像包中的文件夹名字《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》 figureId：人像资源ID，定制人像资源必须传非null值，只有默认人像（目前是10001）可以传null，该值是调用生成人像的AP接口返回的，figureId就是下载的人像包文件夹内的sdk.info文件里的figure值《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》 isCloseAudio：true：关闭声音；false：打开声音（默认） isCloseSilent：true：关闭静默态；false：打开静默态（默认）	人像模型
display Mode	Display Mode	RENDER_ONLY：surface模式	
callback	Change VirtualModelCallback	code：0：成功，其他：失败 subCode：code非0时，可能会有非0的取值 msg：提示消息	code、subCode 参考7.3、错误信息参数说明

调用示例

```
Model model = new Model(useVirtualId);
DisplayMode displayMode = new DisplayMode();
displayMode.setDataMode(DisplayMode.DataMode.RENDER_ONLY); // surface模式
player.setVirtualModel(model, displayMode, (msgType, msgSubType, msg) -> {

});
```

3.1.6、文本转PCM 回调

非必需函数， 如果需要txt相关信息则添加该回调即可

```
boolean setPlayTxtCallback(TxtPlayCallback callback)
```

参数名	类型	取值	说明
callback	TxtPlayCallback	code: 错误码 subCode： txtSn：txt的唯一标志，调用文本播放时自动生成该文本的唯一表示 msg：错误信息	

调用示例

```
player.setPlayTxtCallback(new Player.TxtPlayCallback() {
    @Override
    public void onTtsToPcm(int code, int subCode, String txtSn, String msg) {

    }
});
```

3.1.7、形象模型MD5 校验

非必需函数， 如果校验模型完整性时调用

void virtualModelCheck(Model model, VirtualModelCheckCallback callback);

参数名	类型	取值	说明
model	Model	visualize：形象ID，visualize值默认是10001，内置的默认人像是"慕晴"，如加载定制人像包，visualize值则要大于0且不等于20，也不能等于10001-11000中的数，还不能重复，visualize值为数字人像包中的文件夹名字《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》 figureId：人像资源ID，定制人像资源必须传非null值，只有默认人像（目前是10001）可以传null，该值是调用生成人像的AP接口返回的，figureId就是下载的人像包文件夹内的sdk.info文件里的figure值《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》	人像模型
callback	VirtualModelCheckCallback	code：0：成功，其他：失败 subCode：code非0时，可能会有非0的取值 msg：提示消息	

调用示例

```
Model model = new Model(useVirtualId);
player.virtualModelCheck(model, new Player.VirtualModelCheckCallback() {
    @Override
    public void onVirtualModelCheckStatus(int code, int subCode, String msg) {

    }
});
```

3.2、数字人形象启动 3.2.1、设置显示解码

boolean setDisplay(Surface surface, int width, int height)

参数名	类型	取值	说明
surface	Surface		展示形象的surface
width	int		展示形象的宽
height	int		展示形象的高
返回值	boolean		如果surface为空，返回false

调用示例

```
/**
 * Surface 模式下
 */
@Override
public void surfaceChanged(SurfaceHolder holder, int format, int width, int height) {
    player.setDisplay(holder.getSurface(), width, height);
}
```

3.2.2、形象展示

void onResume()

调用示例

```
@Override
protected void onResume() {
    player.onResume();
    super.onResume();
}
```

3.3、数字人形象暂停

```
void onPause();
```

调用示例

```
@Override
protected void onPause() {
    player.onPause();
    super.onPause();
}
```

3.4、数字人形象切换

注意事项1：Surface 模式在进入人像页面后支持720p切换720p，也支持1080p切换1080p。此外，如果想从720p切换到1080p或从1080p切换到720p，需要重新初始化SDK和重新加载人像。

注意事项2：如果不想使用默认人像，想直接使用定制人像，请先参考本文档的8.1、定制数字人像包替换本地数字人像包《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》，然后重新初始化SDK和重新加载人像（定制人像B替换本地人像A，则需要重新加载一次，定制人像C替换定制人像B也需要重新加载一次），重新加载人像需要在执行demo的MainActivity的initPlayer方法的player.setVirtualModel方法前，初始化定制人像的Model，初始化方法是Model model = new Model(visualize, figureId)，参数说明请参考本文档3.4、数字人形象切换《具体可参考本文档3.4、数字人形象切换》表格的Model说明。

```
void changeVirtualModel(Model model, ChangeVirtualModelCallback callback)
```

参数名	类型	取值	说明
model	Model	visualize：形象ID，visualize值默认是10001，内置的默认人像是"慕晴"，如加载定制人像包，visualize值则要大于0且不等于20，也不能等于10001-11000中的数，还不能重复，visualize值为数字人像包中的文件夹名字《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》 figureId：人像资源ID，定制人像资源必须传非null值，只有默认人像（目前是10001）可以传null，该值是调用生成人像的AP接口返回的，figureId就是下载的人像包文件夹内的sdk.info文件里的figure值《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》	人像模型
callback	ChangeVirtualModelCallback	code：0：成功，其他：失败 subCode：code非0时，可能会有非0的取值 msg：提示消息	code、subCode 参考7.3、错误信息参数说明

调用示例

```
Model model = new Model(useVirtualId);  
// Model model = new Model(useVirtualId, figureId);  
player.changeVirtualModel(model, (code, subCode, msg) -> {  
  
});
```

3.5、数字人形象释放

1、surface模式下释放接口需在view销毁完成后进行调用；例如：有一个surface view，需要在surface view 执行 surfaceDestroyed 回调内进行调用release() 方法

```
void release();
```

调用示例

```
/**  
 * 移除surfaceView  
 */  
ViewGroup parent = (ViewGroup) mSurfaceView.getParent();  
parent.removeView(mSurfaceView);  
mSurfaceView.getHolder().getSurface().release();  
mSurfaceView = null;  
  
/**  
 * surfaceDestroyed 内执行release方法  
 */  
@Override  
public void surfaceDestroyed(SurfaceHolder holder) {  
    player.release();  
    finish();  
}
```

4、普通模式驱动数字人

普通模式下播报数据，与流式模式下不同的是流式模式可持续播报而普通模式单次只可播报一条数据，具体调用参考接口文档

4.1、PCM播报

注意pcm的格式(单声道，16k，16bit)。播放buffer最大只支持10分钟，大约72MB。插播最大3分钟的数据，大约21.9MB；byte[]长度需大于16000，小于16000 x (16 / 8)x 30x20

```
boolean speakWithPcm(Context context, int[][] actions, byte[] bytes);
```

参数名	类型	取值	说明
context	Context		上下文
actions	int[][]	{0, 0} 默认值	播报文本动作，目前不支持文本动作设置
bytes	byte[]		pcm数据， 格式要求： <i>16000hz，单声道，16位</i>
返回值	boolean	true： pcm数据设置成功 false： pcm 数据设置失败	

调用示例

```
byte[] pcmBytes = FileUtils.getBytes(pcmPath + "input/1_16.pcm");
player.speakWithPcm(mContext, new int[][]{{1,10}}, pcmBytes);
```

4.2、文本播报

需要注意的是，当SDK启动后，第一次speak一段文本时，需要请求网络，请保证设备的网络正常。在sdk不重启的情况，如果一段文字请求过，则会缓存在内存中，不再请求网络，这样会占用一部分内存。需要注意pcm的格式(单声道，16k，16bit)。播放buffer最大只支持10分钟，大约72MB。插播最大3分钟的数据，大约21.9MB。默认txt最大为520个字符，对于需要超过这个长度文本，请和百度方沟通。

```
PlayTxtResult speakWithTxt(Context context, String stext, int[][] actions, SpeakerConf conf);
```

参数名	类型	取值	说明
context	Context		上下文
stext	String		播报文本信息。 当使用LITE版克隆音色及公共音色库中支持多语言的音色时，支持自动识别输入语言来进行播报。例：当输入俄文时，使用俄语播报。
conf	SpeakerConf		设置音库， 具体参数请参考 7.2 SpeakerConf 参数
返回值	PlayTxtResult		该文本的唯一标识和执行结果

调用示例

```
player.speakWithTxt(mContext, "你好，我是百度数字人。百度（Baidu）是拥有强大互联网基础的领先AI公司。" +
"百度愿景是：成为最懂用户，并能帮助人们成长的全球顶级高科技公司。“百度”二字，" +
"来自于八百年前南宋词人辛弃疾的一句词：众里寻他千百度。", new int[][]{{1, 10}},
SpeakerConf.builder());
```

4.3、PCM 插播

```
boolean speakInsertOnFlowPcm(Context context, byte[] bytes)
```

参数名	类型	取值	说明
context	Context		上下文
bytes	byte[]		<i>需要插播的pcm数据</i>
返回值	boolean	true： 插播pcm数据设置成功 false：插播pcm 数据设置失败	

调用示例

```
byte[] pcmBytes = FileUtils.getByte(pcmPath + "input/1.pcm");
boolean ret = player.speakInsertOnFlowPcm(mContext, pcmBytes);
```

4.4、文本插播

```
PlayTxtResult speakInsertTxt(Context context, String stext, SpeakerConf conf)
```

参数名	类型	取值	说明
context	Context		上下文
stext	String		需要插播的文本数据
conf	SpeakerConf		设置音库，具体参数请参考 7.2 SpeakerConf 参数
返回值	PlayTxtResult		该文本的唯一标识和执行结果

调用示例

```
PlayTxtResult result = player.speakInsertTxt(mContext,
    "你好啊，虽然我不认识您，但是相逢就是有缘，欢迎来到直播间~老铁~",
    SpeakerConf.builder().setPerson(Config.sTtsPerson));
```

4.5、暂停播报

```
boolean pausePlay()
```

参数名	类型	取值	说明
返回值	boolean	true: 暂停播放 false : 停止异常	如果已经停止的状态，也会返回true

调用示例

```
player.pausePlay();
```

4.6、继续播报

```
boolean resumePlay()
```

参数名	类型	取值	说明
返回值	boolean	true: 暂停播放 false : 停止异常	如果已经停止的状态，也会返回true

调用示例

```
player.resumePlay();
```

4.7、打断播报

打断当前正在进行播报包含插播数据，打断后不可恢复

```
void interrupt()
```

调用示例

```
player.interrupt();
```

4.8、是否正在播放中

用于播放状态判断

```
boolean isPlaying()
```

参数名	类型	取值	说明
返回值	boolean	<i>true: 正在播放中</i> <i>false : 停止播放</i>	

调用示例

```
player.isPlaying()
```

4.9、是否在暂停状态

用于播放状态判断

```
boolean isPause()
```

参数名	类型	取值	说明
返回值	boolean	<i>true: 暂停状态</i> <i>false : 非暂停状态</i>	

调用示例

```
player.isPause()
```

4.10、设置播报音量

按需使用，该能力仅用于音频数据本身较小、存在数学放大条件的前提下可用，若音频原始数据幅值已经过大（最大值参考：-32768~ 32768 ），请使用硬件方式进行功率放大，代码中该能力不可用；
错误信息参考 [3.1.4、添加player消息回调中的onPcmVolumeError回调](#)

```
void setPcmVolume(float vol)
```

参数名	类型	取值	说明
vol	float	默认值为1，范围[1~10]	

调用示例

```
player.setPcmVolume(1);
```

5、流式模式驱动数字人

开启流式播报后可以持续播报数据，与普通模式下不同的是普通模式只能单条播报而流式可持续播报，详细使用可参考本文档4.2.2、流式播报《具体可参考本文档4.2.2、流式播报》，具体调用参考接口文档

需要注意pcm的格式(单声道，16k，16bit)；播放buffer最大只支持10分钟，大约72MB。每次调用speakFlowOnPcm时byte[]的长度需大于16000，开启降噪时需保证降噪后的有效音频byte[]长度大于16000

5.1、开启流式播报 5.1.1、不需要噪音抑制功能

```
boolean speakFlowStart(Context context)
```

参数名	类型	取值	说明
context	Context		上下文
返回值	boolean	<i>true: 开启成功</i> <i>false : 开启失败</i>	

调用示例

```
player.speakFlowStart(mContext);
```

5.1.2、需要噪音抑制功能

```
boolean speakFlowStart(Context context, boolean startNoiseSuppressor)
```

参数名	类型	取值	说明
context	Context		上下文
startNoiseSuppressor	boolean	<i>true : 开启噪音抑制</i> <i>false : 不开启噪音抑制</i>	<i>是否开启噪音抑制</i>
返回值	boolean	<i>true: 暂停状态</i> <i>false : 非暂停状态</i>	

调用示例

```
player.speakFlowStart(mContext, true);
```

5.2、关闭流式播报

```
boolean speakFlowFinish(Context context)
```

参数名	类型	取值	说明
context	Context		上下文
返回值	boolean	<i>true: 暂停状态</i> <i>false : 非暂停状态</i>	

调用示例

```
player.speakFlowFinish(mContext);
```

5.3、流式PCM播报

```
boolean speakFlowOnPcm(Context context, byte[] bytes)
```

参数名	类型	取值	说明
context	Context		上下文
bytes	byte[]		需要插播的pcm数据
返回值	boolean	true : PCM 数据设置成功 false : PCM 数据设置失败	

调用示例

```
byte[] pcmBytes = FileUtils.getBytes(pcmPath + "input/1_16.pcm");
player.speakFlowOnPcm(mContext, pcmBytes);
```

5.4、流式TXT播报

```
PlayTxtResult speakFlowOnSegTxt(Context context, String segTxt, SpeakerConf conf)
```

参数名	类型	取值	说明
context	Context		上下文
segTxt	String		需要插播的文本数据
conf	SpeakerConf		设置音库，具体参数请参考 7.2 SpeakerConf 参数
返回值	PlayTxtResult		该文本的唯一标识和执行结果

调用示例

```
PlayTxtResult result = player.speakFlowOnSegTxt(mContext, "你好，我是百度数字人。百度（Baidu）是拥有强大互联网基础的领先AI公司。百度愿景是："+
    "成为最懂用户，并能帮助人们成长的全球顶级高科技公司。“百度”二字，来自于八百年前南宋词人辛弃疾的一句词："+
    "众里寻他千百度。",SpeakerConf.builder().setPerson(Config.sTtsPerson));
```

5.5、开始流式插播

开启插播后进入插播模式，需要调用关闭插播接口进行退出插播模式

```
boolean speakInsertFlowPcmStart(Context context)
```

参数名	类型	取值	说明
context	Context		上下文
返回值	boolean	true: 开启插播正常 false : 开启插播异常，可能未处于流式播报模式	

调用示例

```
player.speakInsertFlowPcmStart(mContext)
```

5.6、关闭流式插播

```
boolean speakInsertFlowPcmFinish(Context context, boolean isFnishInsertPcm)
```

参数名	类型	取值	说明
context	Context		上下文
isFnishInsertPcm	boolean	<i>true: 播放完之前已经插入的数据</i> <i>false: 直接停止，回到插播前的内容上（默认）</i>	是否停止插入PCM数据
返回值	boolean	<i>true: 关闭插播正常</i> <i>false：关闭插播异常，可能未开启插播模式</i>	

调用示例

```
player.speakInsertFlowPcmFinish(mContext, true);
```

5.7、流式PCM插播

- 1、如果未开启插播模式，可以支持插播，插播完毕后继续播放正文；如果开启插播则处于插播模式当中，需要调用关闭插播接口，退出插播模式后继续播报正文

2、需要注意pcm的格式(单声道16k; 16bit)。插播最大3分钟的数据，大约21.9MB。

```
boolean speakInsertOnFlowPcm(Context context, byte[] bytes)
```

参数名	类型	取值	说明
context	Context		上下文
bytes	byte[]		<i>需要插播的pcm数据</i>
返回值	boolean	true：插播pcm数据设置成功 false：插播pcm 数据设置失败	

调用示例

```
byte[] pcmBytes = FileUtils.getBytes(pcmPath + "input/1.pcm");
boolean ret = player.speakInsertOnFlowPcm(mContext, pcmBytes);
```

5.8、流式TXT插播

- 1、如果未开启插播模式，可以支持插播，插播完毕后继续播放正文；如果开启插播则处于插播模式当中，需要调用关闭插播接口，退出插播模式后继续播报正文

```
PlayTxtResult speakInsertTxt(Context context, String stext, SpeakerConf conf)
```

参数名	类型	取值	说明
context	Context		上下文
stext	String		需要插播的文本数据
conf	SpeakerConf		设置音库，具体参数请参考 7.2 SpeakerConf 参数
返回值	PlayTxtResult		该文本的唯一标识和执行结果

调用示例

```
PlayTxtResult result = player.speakInsertTxt(mContext,
    "你好啊，虽然我不认识您，但是相逢就是有缘，欢迎来到直播间~老铁~",
    SpeakerConf.builder().setPerson(Config.sTtsPerson));
```

5.9、暂停播报

```
boolean pausePlay()
```

参数名	类型	取值	说明
返回值	boolean	true: 暂停播放 false : 停止异常	如果已经停止的状态，也会返回true

调用示例

```
player.pausePlay();
```

5.10、继续播报

```
boolean resumePlay()
```

参数名	类型	取值	说明
返回值	boolean	true: 暂停播放 false : 停止异常	如果已经停止的状态，也会返回true

调用示例

```
player.resumePlay();
```

5.11、打断播报

打断当前正在进行播报包含插播数据，打断后不可恢复

```
void interrupt()
```

调用示例

```
player.interrupt()
```

5.12、是否正在播放中

用于播放状态判断

```
boolean isPlaying()
```

参数名	类型	取值	说明
返回值	boolean	<i>true: 正在播放中</i> <i>false : 停止播放</i>	

调用示例

```
player.isPlaying()
```

5.13、是否在暂停状态

用于播放状态判断

```
boolean isPause()
```

参数名	类型	取值	说明
返回值	boolean	<i>true: 暂停状态</i> <i>false : 非暂停状态</i>	

调用示例

```
player.isPause()
```

5.14、设置播报音量

按需使用，该能力仅用于音频数据本身较小、存在数学放大条件的前提下可用，若音频原始数据幅值已经过大（最大值参考：-32768 ~ 32768 ），请使用硬件方式进行功率放大，代码中该能力不可用；
错误信息参考 [3.1.4、添加player消息回调中的onPcmVolumeError回调](#)

```
void setPcmVolume(float vol)
```

参数名	类型	取值	说明
vol	float	默认值为1，范围[1~10]	

调用示例

```
player.setPcmVolume(1);
```

6、文本处理

player获取方式参考 [3.1、数字人形象初始化](#)

6.1、预加载文本的pcm数据

boolean preLoadTxtPcm(String segTxt, SpeakerConf conf, TtsPcmDataCallback dataCallback)

参数名	类型	取值	说明
segTxt	String		文本信息
conf	SpeakerConf		设置音库, 详情参数请参考 7.2、 SpeakerConf 参数
callback	TtsPcmDataCallback	code : 0 : 成功, 其他 : 失败 subCode : code非0时, 可能会有非0的取值 msg : 提示消息 bytes : pcm数据	文本转PCM后的回调 code、subCode 参考 7.3、错误信息参数说明

调用示例

```
player.preLoadTxtPcm(txt, SpeakerConf.builder().setPerson(Config.sTtsPerson), new Player.TtsPcmDataCallback() {  
    @Override  
    public void onTtsPcmData(int code, int subCode, String msg, byte[] bytes) {  
  
    }  
});
```

6.2、设置预加载pcm的缓存空间的大小

void setPreLoadPcmCacheSize(int size);

参数名	类型	取值	说明
size	int	取值范围【0 ~ 2048】，默认：1024	单位：MB

调用示例

```
player.setPreLoadPcmCacheSize(1024);
```

6.3、删除文本预加载时缓存的所有文件

void clearCache();

调用示例如下：

```
TtsCacher.instance().clearCache();
```

6.4、删除文本预加载时缓存的某一个文件

void clearTxtCache(String txt, SpeakerConf conf) ;

参数名	类型	取值	说明
txt	String		需要删除的文本信息
conf	SpeakerConf		设置音库, 详情参数请参考 7.2、 SpeakerConf 参数

调用示例如下：


```
TtsCacher.instance().clearTxtCache("需要删除的文本信息", SpeakerConf.builder().setPerson(Config.sTtsPerson));
```

7、工具类 7.1、模型效验（包括文件完整性校验和文件版本校验）

```
void checkModel(Model model, ModelCheck.ModelCheckCallback callback);
```

参数名	类型	取值	说明
model	Model	visualize：形象ID，visualize值默认是10001，内置的默认人像是"慕晴"，如加载定制人像包，visualize值则要大于0且不等于20，也不能等于10001-11000中的数，还不能重复，visualize值为数字人像包中的文件夹名字《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》 figureId：人像资源ID，定制人像资源必须传非null值，只有默认人像（目前是10001）可以传null，该值是调用生成人像的API接口返回的，figureId就是下载的人像包文件夹内的sdk.info文件里的figure值《具体可参考本文档8.1、定制数字人像包替换本地数字人像包》	人像模型
callback	ModelCheck.ModelCheckCallback	figureId：人像资源id，调用生成人像的API接口返回的。内置的默认人像是"慕晴"，传null code：参考7.3、错误信息参数说明 subCode: 参考7.3、错误信息参数说明 msg：提示消息	

code取值

取值	说明
0	校验成功
-1000	检查失败，目录不存在
-1001	检查失败，版本信息文件不存在
-1002	检查失败，版本信息文件内容为空
-1003	检查失败，版本信息文件内容格式错误
-1004	检查失败，人像资源包版本过低
-1005	检查失败，默认模型init文件不存在
-1006	检查失败，后台返回为空
-1007	检查失败，后台状态不成功
-1008	检查失败，没有返回文件信息
-1009	检查失败，一般是网络异常
-1010	检查失败，模型文件个数为0
-1011	检查失败，init文件内容为空
-1012	检查失败，init文件中的文件个数为0
-1013	检查失败，init文件格式错误
-1014	检查失败，文件不匹配，文件缺失或被修改
-1015	检查失败，文件匹配对比时异常

调用示例

```
VirtualModelUtil.checkModel(model, (figureId, code, subCode, msg) -> {  
  
});
```

7.2、SpeakerConf 参数

需要注意的是，需要和百度方确认，per取值的范围

参数名	类型	取值	说明
spd	int	语速，取值0-15，默认为5中语速	语速
pit	int	音调，取值0-15，默认为5中音调, 克隆音色暂不生效	音调
vol	int	音量，取值0-15，默认为5中音量，克隆音色暂不生效	音量
aue	int	6：wav（内容同pcm-16k）默认	音频格式
per	int	默认：度小希=5116 更多音色请参考： https://cloud.baidu.com/doc/AI_DH/s/Slywt3fxy	音色

7.3、错误信息参数说明

code	subCode	参数含义
0	0	成功
-1000	-2000	参数错误
	-2006	获取tts异常
	-2007	tts 返回null
-1001	-2001	重复设置模型
	-2003	调用顺序异常
	-2005	正在播放中
	-2009	没有播报
-1002	-2002	形象包不存在
	-2008	缓存已满
-1003	-2008	插播缓存已满
-1004	-2004	数字人切换频繁

8、常见问题 8.1、定制人像包替换本地数字人像包 SDK支持定制人像包替换

- 命名定制人像包：

如使用定制人像包，则在初始化SDK之前需要下载对应的定制人像包压缩文件，定制人像包压缩文件解压后就是一个人像包文件夹，人像包文件夹需要改名字，用数字来命名（数字大于0且不等于20，也不能等于10001-11000中的数，还不能重复）。

- push路径：

把人像包文件夹push到设备的

/sdcard/baidu/aip/virtual_models/data/render_model路径内（此路径是定制人像包的路径，使用adb push命令把人像包文件夹push进去）。

- 使用方式：

定制人像包的路径会多一个文件夹，然后调用切换数字人方法可切换到此定制的数字人形象《具体可参考本文档3.4、数字人形象切换》。

3D数字人交互SDK-H5版

1. 使用流程

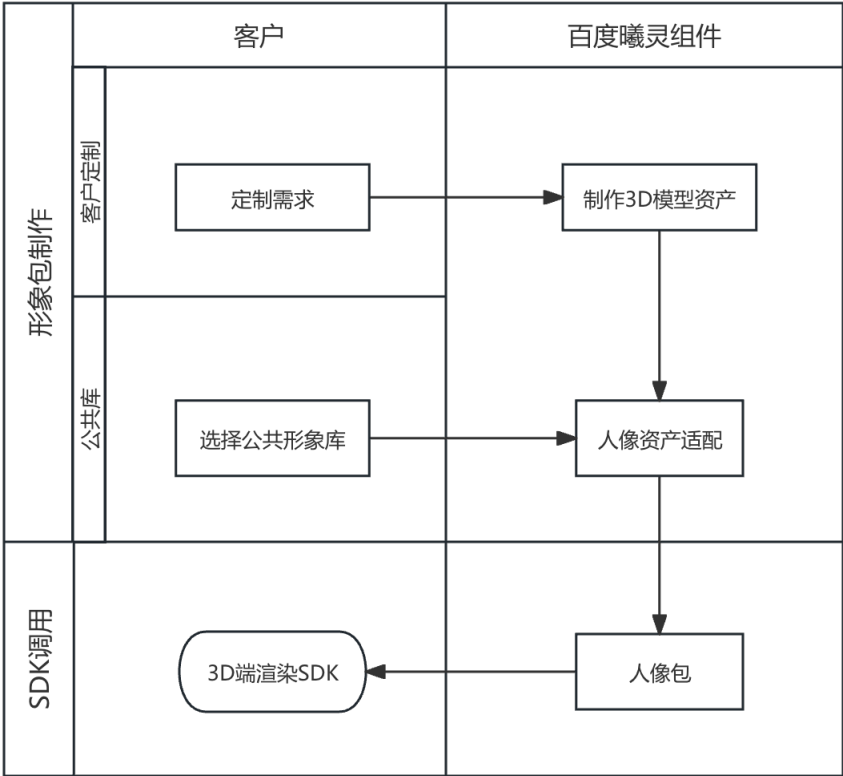
1、形象包制作：分为“定制形象”和“公共库形象”两种。

定制形象

用户提供设计原画，百度侧根据需求制作3D人物模型，为保证人物模型可以被端渲染SDK驱动使用，3D端渲染人像需要进行适配。预计需要3-4周时间。人像制作过程中允许用户调整和修改，确保人像符合用户需求。

公共形象库

参考文档底部《附.公共人像库》



形象定制及适配费用相关问题，可咨询：[咨询单](#)



2、SDK调用：

客户端或页面集成SDK加载形象包，本地驱动渲染数字人。

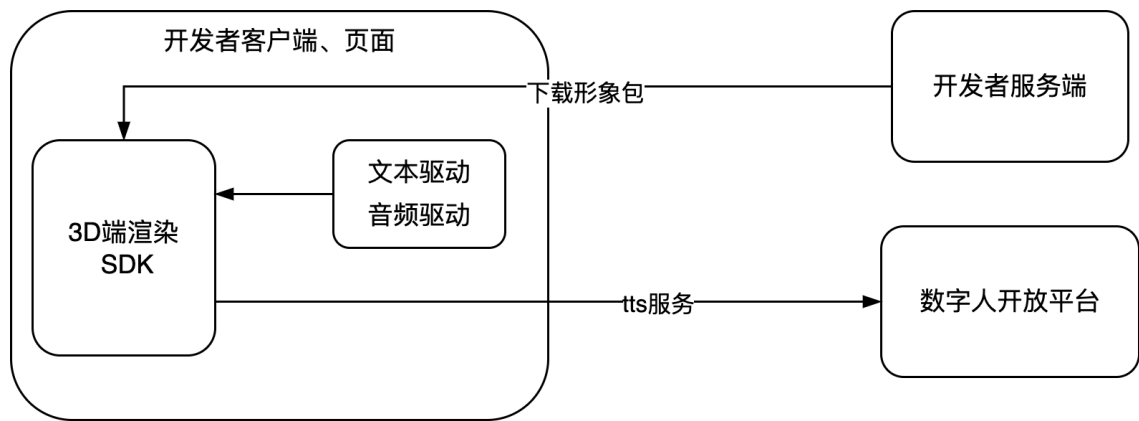
支持音频和文本驱动:

音频可以使用真人音频或自行接入的第三方的TTS音频。

文本驱动SDK内部会调用数字人开放平台TTS服务驱动数字人，

文本驱动可以开启智能动作开关，实现数字人根据播报语义智能触发动作。

3D数字人形象包可切换机位，动作驱动，切换配饰服装（如有多套配置服装）



购买入口：[组件商店](#)（依次购买：1. 定制服务-语音合成 2. 交互组件-端渲染交互组件-3D数字人）

2. 接入方法

iframe支持在网页内引入其他网页，使用iframe引入WebSdk方式如下：

FAQ: 为什么使用iframe

使用iframe的方式天然的和集成方代码形成隔离，形成单独的进程进行数字人的渲染工作，防范因js单线程设计可能造成的渲染卡顿问题。

```
<iframe
  id="myIframe"
  src="https://open.xiling.baidu.com/cloud/client/ar?platform=paas&licenseKey=自行填写&licenseUrl=自行填写"
  allow="autoplay;"
  style="height: 400px;width: 400px;"
/>
```

2.1 url query参数说明（均为必填） | 参数 | 解释 | ----- | platform | 请填写为：paas | | licenseUrl | 联系客服，申请的鉴权文件网络地址（不可为file://开头的本地地址） | | licenseKey | 联系客服，申请的鉴权文件的文件名称 |

3. demo 使用

3.1 demo说明

此demo分别使用react，vue两种框架。

针对如何初始化数字人，发送消息驱动数字人，发送流式音频驱动数字人，以及数字人播报的打断，数字人的销毁，动作驱动，机位切换等功能给予了相应的实现。

3.2 demo下载

React版本，下载地址<https://sdk-demo.bj.bcebos.com/client-human-react-demo.zip>

Vue版本，下载地址<https://sdk-demo.bj.bcebos.com/client-human-vue-demo.zip>

3.3 demo运行

- 判断环境是否有 node(22 版本以上)和 yarn，如没有则进行 node, yarn 的安装： brew install node; npm install --global yarn (其他安装方式请自行搜索)
- 解压后，进入目录：
 - React版本目录名称：client-human-react-demo

- Vue 版本目录名称：client-human-vue-demo
1. 命令行执行yarn，安装所需依赖包（根据提示使用对应的 node 版本）
 2. 搜索“自行填写”关键字，将demo中一些需要自己填写的，如形象包地址、token等变量进行填写
 3. 文本和音频驱动的token需要自行生成，注：从应用管理中获取appKey/appld后, 参见[token生成文档](#)，生成token。
 4. 命令行执行npm run dev，启动项目
 5. 将控制台输出的ip地址，复制到浏览器地址栏进行项目预览

3.4 SDK兼容性 支持机型信息：

- 安卓：支持CPU骁龙730G/骁龙855/联发科天玑900/天玑820 同等性能及以上芯片
- iOS：支持iPhone8 系统版本13.3以及以上。

支持浏览器类型：

机型	系统浏览器	qq浏览器	百度浏览器	微信内置浏览器	UC浏览器
华为mate 50 (鸿蒙系统)	测试版本：15.0.8.301	测试版本：15.7.0.0081 64bit	测试版本：13.74.0.10	测试版本：8.0.53	测试版本：V17.2.0.1351
小米 11pro	测试版本：18.6.71011	测试版本：15.7.2.2041	测试版本：13.74.0.10	测试版本：8.0.53	测试版本：17.1.9.1350
oppo reno6 pro	测试版本：40.8.29_db2babb	测试版本：14.9.1.1.1033 64bit	测试版本：13.53.0.10	测试版本：8.0.49	测试版本：V16.3.9.1290
iPhone15pro max	测试版本：ios17.5.1	—	测试版本：13.73.1.10	测试版本：8.0.53	测试版本：17.2.0.2476
iPhone12	测试版本：ios17.6.1	—	测试版本：13.53.5.10	测试版本：8.0.48	测试版本：17.2.0.2476
iphone xs	测试版本：ios15.3	—	测试版本：13.74.1.10	测试版本：8.0.53	测试版本：17.1.9.2473
iphone 14pro	测试版本：ios16.3.1	—	测试版本：13.74.1.10	测试版本：8.0.54	测试版本：17.1.9.2476

(其余浏览器及上述浏览器低版本可能存在不兼容情况)

4. SDK 的使用

4.1 SDK 驱动原理及工具包的使用

4.1.1 SDK 驱动原理

SDK 的驱动原理是通过业务页面与数字人iframe之间的消息发送和状态监听来实现数字人的初始化，以及数字人状态的监听；

这里我们封装了一个 npm 包来便捷客户进行上述消息的发生和监听，用户只需调用 npm 包提供的方法即可完成数字人的初始化和消息的发送以及回调的监听。

4.1.2 工具包的安装

```
// npm 方式 :
npm i @bddh/starling-dhiframe
// yarn 方式
yarn add @bddh/starling-dhiframe
```

4.1.3 工具包的使用

```
import {DHlframeV2} from '@bddh/starling-dhiframe';

const dhlframe = new DHlframeV2('iframeId'); // 这里修改成 iframe 的 id属性值
// 模型的初始化
dhlframe.load(...)
// 文本驱动和回调
dhlframe.textRender(...)
```

4.2 SDK初始化以及状态监听

4.2.1 监听 iframe引擎加载完成

```
dhlframe.registerMessageReceived((data: any) => {
  if (data.type === 'state') {
    if (data.content.status === 'INIT') {
      console.log('引擎初始化完成')
      dhlframe.load({
        backgroundImageUrl: 'https://digital-human-js-
cdn.cdn.bcebos.com/meta_human_editor/20240903202457/digital-human-open-console-web/img/bg1.png',
        modelUrl: '定制人像包的zip的cdn地址'
      });
    }
    if (data.content.status === 'AVATAR_LOAD') {
      console.log('模型数字人加载完成')
    }
    if (data.content.status === 'CLOSE') {
      console.log('模型数字人销毁完成')
    }
  }
});
```

4.2.2 status 状态列表

data.content.stat us	含义	备注
INIT	引擎初始化完成	引擎初始化完成之后，才可以使用dhlframe.load方法加载人像包
ERROR	启动发生错误	INIT之前发生ERROR是引擎发生错误INIT之后传入人像包发生错误是人像包加载失败
AVATAR_LOAD	模型包数字人加载完成	dhlframe.load完成的状态提示
USER_ACTIVE	数字人加载时无用户点击交互	场景：使用开场白驱动数字人时，通过监听该消息，进行用户交互提醒
CLOSE	模型包数字人销毁完成	dhlframe.destroy完成后的状态提示
AUTHENTICATION	鉴权回调信息	{ code: '错误码' message: '错误信息'}

鉴权失败code码

错误码	错误信息
1001	服务器内部错误
1002	APP已失效
1003	APP token已过期
1004	目前线上用户较多，请您稍后再试或tts字符包配额不足
1005	授权信息不存在
1008	App 签名非法
1009	App key非法
1010	App token非法
1012	参数不能为空
1202	正在连接或已连接
1204	解析请求失败
1206	读取json失败

4.2.3 加载人像包

```
dhlframe.load({
  backgroundImageUrl: 'https://digital-human-js-cdn.cdn.bcebos.com/meta_human_editor/20240903202457/digital-human-open-console-web/img/bg1.png',
  modelUrl: '定制人像包的zip的文件地址',
  cameraId: 'face_cam'
});
```

参数	必填	含义
modelUrl	是	模型包文件地址，不支持本地文件地址；开发者向运营人员获取所需人像zip包文件，存储到自有文件服务后填写文件下载 URL。建议使用 CDN 地址以提高加载速度。
backgroundImageUrl	否	人像背景，不填为透明背景，不支持本地文件地址，使用bos或cdn获取url
cameraId	否	加载完成人像资源后，人像的初始相机位（全身，半身等），参见人像zip包中config.json文件

tip: load方法需要在监听到status为INIT后才可以进行调用

4.3 文本驱动

```
dhlframe.textRender({
  // 购买"定制服务-语音合成 交互组件-端渲染交互组件-3D数字"两个组件组件，为应用绑定组件，获取appKey, appId
  // 参见token生产文档（https://cloud.baidu.com/doc/AI_DH/s/Ulywupd35）生成token
  token: '填写 appKey, appId 生成的鉴权 token',
  text: '使用文本驱动数字人的播报文本',
  tts: {
    per: 5116,
    spd: 5,
    pit: 5
  },
},
({status, data}) => {
  Logger.log('textRender callback:', status, data);
});
);
```

参数	必填	类型	含义
text	是	string	播报文本, 支持 在线SSML（语音合成标记语言） ，不可包含& < >特殊字符 •当使用LITE版克隆音色及公共音色库中支持多语言的音色时，支持自动识别输入语言来进行播报。例：当输入俄文时，使用俄语播报。
token	是	string	鉴权 token，生产方式参见： 鉴权参数生成指南
tts	是	{ per: number string; spd: number; pit: number; }	音色发音人ID，可用发音人列表参考： 公共音色库 ，或使用[克隆音色] (https://cloud.baidu.com/doc/AI_DH/s/qm1ftwtgi spd) 语速 5是正常值，0-15的取值范围。越大语速越快，默认值5（字面量需要是整数） pit 语调 5是正常值，0-15的取值范围，越大声音越尖，默认值5（字面量需要是整数）
characterImage	否	string	人像id，用于标识人像，从人像包 config.json 中获取。目前只用于开启智能动作功能，参考下面 autoAnimoji 字段说明。
autoAnimoji	否	boolean	是否为本次文本驱动智能匹配动作， 当填写有效 characterImage 且本参数设为 true 时开启。
listener	否	{status, data}	staus为驱动回调的状态，staus枚举及含义如下

listener中staus枚举及说明

status	含义	data
RENDER_START	文本驱动开始的状态回调	-
RENDER_COMPLETED	文本驱动完成的状态回调	-
RENDER_FINISH	驱动完成或被打断状态的回调如被打断，则状态从RENDER_START直接变为RENDER_FINISH不存在RENDER_COMPLETED	-
RENDER_WARNING	缺少用户点击交互行为，无法有声播放	{code: 4000}
SUBTITLE	播报内容	{content: '当前播报的文本'}
RENDER_ERROR	文本驱动错误	{code: number}错误码描述如下表

RENDER_ERROR中错误码映射表如下：

code错误码	错误信息	说明
1604	characterImage非法	characterImage填写错误
1203	渲染文本失败	body参数非法，不支持 body 传& < >等特殊字符
1208	富文本格式非法	SSML格式错误
1201	操作尚未支持	--
23002	ttsPer非法	tts ttsPer参数填写非法
4001	文本不能为空	--

token鉴权失败消息回调4.2.2 status中AUTHENTICATION描述

4.4 音频驱动


```
// 购买"定制服务-语音合成 交互组件-端渲染交互组件-3D数字"两个组件组件，为应用绑定组件，获取appKey, appld
// 参见token生产文档（https://cloud.baidu.com/doc/AI_DH/s/Ulywupd35）生成token
dhlframe.startAudioRender('填写appKey,appld生成的token', ({status, data}) => console.log('驱动回调消息：', status, data));
dhlframe.audioRender(data); //ArrayBuffer音频数据(PCM 编码，16k, 单通道，16bit)
dhlframe.stopAudioRender(); //结束音频渲染
```

驱动方法	传参	备注
startAudioRender	(token: string,listener?: (msg) => void)	开启语音驱动并设置状态回调函数 msg枚举 详见下表
audioRender	(data: ArrayBuffer)	传输音频数据(PCM 编码，16k, 单通道，16bit)
stopAudioRender	-	结束语音驱动

listener中**msg枚举**及说明

status	含义	data
AUDIO_INIT	音频驱动初始化完成，准备接收数据的状态回调	-
RENDER_COMPLETED	音频驱动完成的状态回调	-
RENDER_FINISH	驱动完成，如被打断，则状态从RENDER_START直接变为RENDER_FINISH不存在RENDER_COMPLETED	-
RENDER_WARNING	缺少用户点击交互行为，无法有声播放	code码为4000
RENDER_ERROR	音频驱动错误	errMsg信息

4.5 其他驱动指令

```
// 销毁数字人
dhlframe.destroy();
// 打断数字人播报
dhlframe.interrupt();
```

驱动指令	传参	备注
interrupt	-	打断当前播报内容
destroy	-	销毁数字人
changeCamera	(camerald: string;)	切换机位，参见zip包中config.json文件
animojiRender	(animojild: string, listener?: (msg) => void)	切换动作，参见zip包中config.json文件
changeParts	(parts: Array({ id: string; // 资产id type?: string; //资产类型}))	切换数字人相关配饰，参见zip包中config.json文件例：dhlframe.changeParts([{ type: "hair", id: "coco_02_hair_out" }, { type: "badge", id: "coco_02_badge_out" }]);

5.常见问题

5.1 本地调试过程中数字人发生形变

数字人销毁，需要通过指令进行，本地热更新时可能触发数字人模型非正常销毁的情况

解决方法：1. 刷新浏览器；2. 正确调用dhiframe.destroy进行销毁数字人

5.2 控制数字人播放音量

```
import {DHlframeV2, MESSAGE_TYPE_V2} from '@bddh/starling-dhiframe';

// vol为0，则为静音
dhlframe?.sendMessage(MESSAGE_TYPE_V2.COMMAND, {vol: 0});
// vol为1，则为恢复正常音量
dhlframe?.sendMessage(MESSAGE_TYPE_V2.COMMAND, {vol: 1});
```

5.3 点击iframe内元素的回调事件

```
dhiframe.registerMessageReceived((data: any) => {
  if (data.type === 'CLICK-HUMAN') {
    console.log('点击了iframe内部')
  }
})
```

5.4 数字人播报只有口型没有声音 根因：浏览器在无用户交互（交互：用户点击，拖拽等）的情况下，播放有声音视频，会自动开启静音；

解决方案：在数字人播报之前，引导用户进行点击等行为操作。（部分浏览器点击行为的作用目标为iframe）

具体步骤示例：

1. iframe上层添加div，为div设置pointer-events: none;
2. 根据5.3监听iframe点击回调，进行div元素的隐藏
3. 点击交互完成，进行后续数字人驱动流程

5.5 js 和 license跨域cors报错 获取到的 zip 包和 license在上传到cdn存储后，需要设置cors白名单；

提醒：强烈建议使用cdn存储，使用oss或服务器存储，需要配置response header以及开启gzip压缩提速等

3D数字人交互SDK-Android版

简介

本文说明Android 端渲染SDK的使用和调用方式

SDK下载

1.0.6 版本 [下载地址](#)

SDK包体大小：33mb（不包含人像包大小）

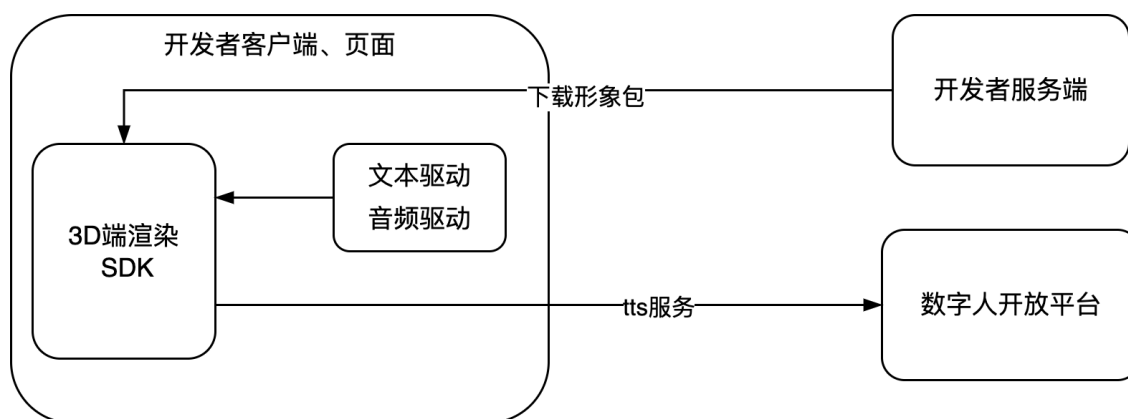
支持机型信息：

- 安卓：支持CPU骁龙730G/骁龙855/联发科天玑900/天玑820 同等性能及以上芯片
- iOS：支持iPhone8 系统版本13.3以及以上。

整体架构&功能简介

- 客户端或页面集成SDK加载形象包，本地驱动渲染数字人。
- 支持音频和文本驱动:
- 音频可以使用真人音频或自行接入的第三方的TTS音频。
- 文本驱动SDK内部会调用数字人开放平台TTS服务驱动数字人，

- 文本驱动可以开启智能动作开关，实现数字人根据播报语义智能触发动作。
- 3D数字人形象包可切换机位，动作驱动，切换配饰服装（如有多套配置服装）



请与百度运营人员联系，或者测试鉴权信息和人像包！

鉴权&申请

license需要额外申请，提供给百度运营人员如下信息，用于生成license文件

1. 包名
2. 签名文件的md5值

license文件获取后，必须放在assets目录下，

注意license的有效期，在license失效前，必须强制更新带有新license的app

```
Avatar3DConfig avatar3DConfig = digitalHumanSettings.getAvatarConfig();
// asset 目录下的xuniren-avatar-debug-0004.license文件
avatar3DConfig.setLicenseFilename("xuniren-avatar-debug-0004");
```

包名和签名md5获取

包名：每个 Android 应用均有一个唯一的应用 ID。安卓以 Java 包的形式管理应用。

demo的包名定义在app/build.gradle中

```
applicationId "com.baidu.ar.avatarsdksample"
```

具体说明见官方文档：

<https://developer.android.com/studio/build/application-id.html>

签名md5

生成签名文件：

<https://developer.android.com/studio/publish/app-signing.html>

方法1：

```
命令行输入 keytool -list -v -keystore <签名文件路径>
```

方法2：

使用demo中的签名文件和自己的包名，报错中有md5值

```
DemoAvatarCallback onError: name=license.code=2301,subCode=13,message=onVirtualHumanAuth failed :13
licenseKey:xuniren-avatar-debug-0004,algorithmId:32,packageName:com.baidu.aip.digitalhuman.avatar,md5:6B92C246956
n
```

DEMO测试 & 资源准备

1. 从运营人员处获取license鉴权文件，按照上文提供包名和鉴权文件的md5。demo的包名是 com.baidu.aip.digitalhuman.avatar，签名文件的md5是 6B92C246956520E715F4078D2A6E2961。复制到assets目录下，并填写文件名到demo中
2. 从运营人员处获取人像资源包
3. 从[组件商店](#) 购买《交互组件-端渲染交互组件-3D数字人》及《定制服务-语音合成》。
4. 选择一个音色[公共音色库](#)或者[克隆音色](#)

购买入口：[组件商店](#)（依次购买：1. 定制服务-语音合成 2. 交互组件-端渲染交互组件-3D数字人）

具体DEMO测试代码修改

使用较新的android studio导入即可运行， gradle 的版本为8.8，需要jre 17或以上版本运行

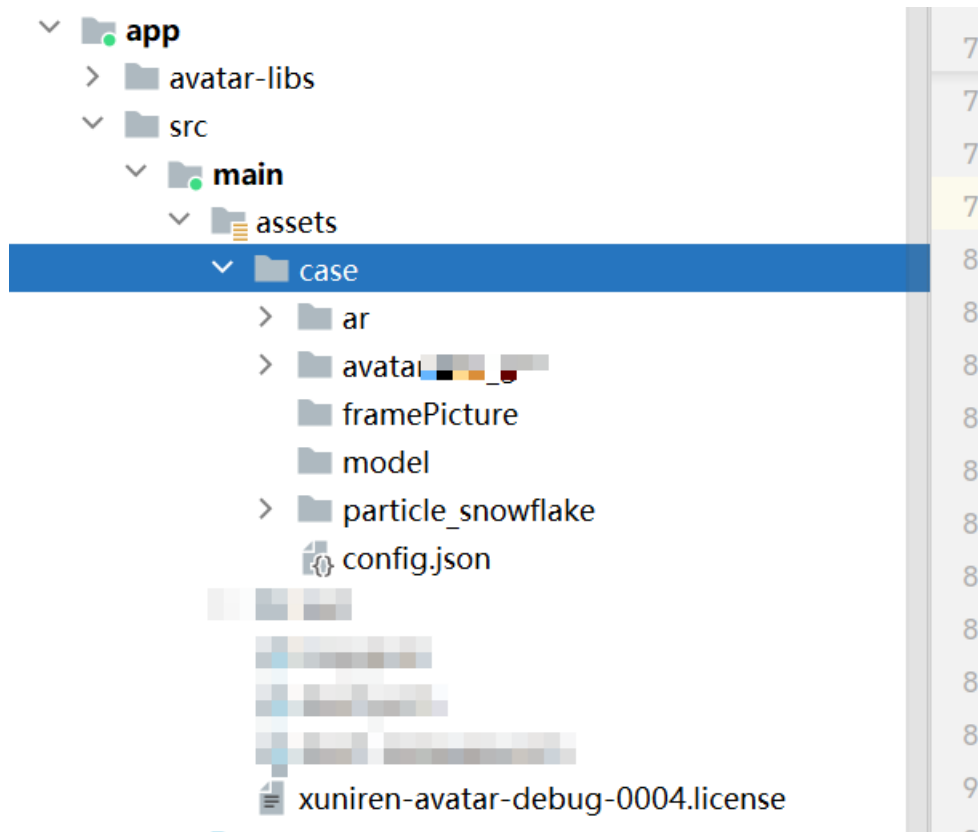
从运营人员处获取license鉴权文件和人像资源包

1. 申请license文件(见下文的[鉴权&申请](#))，需要你的包名和签名文件的md5，获取后的license文件放入到assets目录下
2. 修改 avatar3DConfig.setLicenseFilename("xuniren-avatar-debug-0004"); 里面的"xuniren-avatar-debug-0004"为xuniren-avatar-debug-0004.license为你的license文件名

注意license的有效期限，在license失效前，必须强制更新带有新license的app

```
Avatar3DConfig avatar3DConfig = digitalHumanSettings.getAvatarConfig();
// asset 目录下的xuniren-avatar-debug-0004.license文件
avatar3DConfig.setLicenseFilename("xuniren-avatar-debug-0004");
```

3. 从运营人员获取人像包。在assets目录下，解压人像资源包到case目录下，assets目录结构（含license）如下



从 [组件商店](#) 购买《交互组件-端渲染交互组件-3D数字人》及《交互组件-语音合成》，购买后在「应用管理」创建应用并绑定组件，来获取AppId和AppKey

AppId和AppKey的获取请参照https://cloud.baidu.com/doc/AI_DH/s/plyy6xhi0，注意数字人文字驱动需要额外绑定「语音合成」组件。

1. demo启动后，选择如下3种方式的1种

- 在代码里修改appid、appkey后运行demo
- demo运行后，可以在UI界面中填写 appid， appKey（token无需填写）
- demo运行后，可以在UI界面中填写token（token填写后，appid， appkey无效）

2. tts发音人音色可以使用[公共音色库](#)或者[克隆音色](#)

```
/**
 * 网页上的应用鉴权信息
 */
public static final String DEFAULT_APP_ID = "i-qjwqyvdn3e40y";

/**
 * 网页上的应用鉴权信息
 */
public static final String DEFAULT_APP_KEY = "n5j0cyq6jnnkh8r573f5";

/**
 * tts音色
 */
public static final int TTS_PERSON = 5116;

/**
 * 设置的token, 如果为空，则用appid appkey计算， 建议服务端生成
 */
private static final String DEFAULT_TOKEN = "";
```

demo初始化运行异常排查思路

1. 进入第二个Activity后，有背景图，无人像，点按钮app退出
判断为license错误
 - assets目录下确保有xuniren-avatar-debug-0004.license
 - 修改 `avatar3DConfig.setLicenseFilename("xuniren-avatar-debug-0004");` 里面的"xuniren-avatar-debug-0004"为xuniren-avatar-debug-0004.license为你的license文件名
2. 有数字人，动作正常，但是点击RENDER按钮后，无反应
确认appId，appkey是否填对，网页上勾选"3D端渲染"+"语音合成"权限
如果"播放音频"有声音，那么缺"语音合成"权限

正式鉴权&申请

license需要额外申请，提供给百度运营人员如下信息，用于生成license文件

1. 包名
2. 签名文件的md5值

license文件获取后，必须放在assets目录下

注意license的有效期，在license失效前，必须强制更新带有新license的app

```
Avatar3DConfig avatar3DConfig = digitalHumanSettings.getAvatarConfig();  
// asset 目录下的xuniren-avatar-debug-0004.license文件  
avatar3DConfig.setLicenseFilename("xuniren-avatar-debug-0004");
```

包名和签名md5获取

包名：每个 Android 应用均有一个唯一的应用 ID。安卓以 Java 包的形式管理应用。

```
// demo的包名定义在app/build.gradle中  
  
applicationId "com.baidu.ar.avatarsdksample"
```

具体说明见官方文档：<https://developer.android.com/studio/build/application-id.html>

签名md5

生成签名文件：<https://developer.android.com/studio/publish/app-signing.html>

使用demo中的签名文件和自己的包名，报错中有md5值

```
DemoAvatarCallback onError: name=license,code=2301,subCode=13,message=onVirtualHumanAuth failed :13  
licenseKey:xuniren-avatar-debug-  
0004,algorithmId:32;packageName=com.baidu.aip.digitalhuman.avatar,md5:6B92C246956520E715F4078D2A6E2961;dev-  
-isolated;licenseSdkVersion:5.0.0.9-isolated,exceptionClassName=AvatarAuthException
```

形象定制（使用公共形象可以跳过）

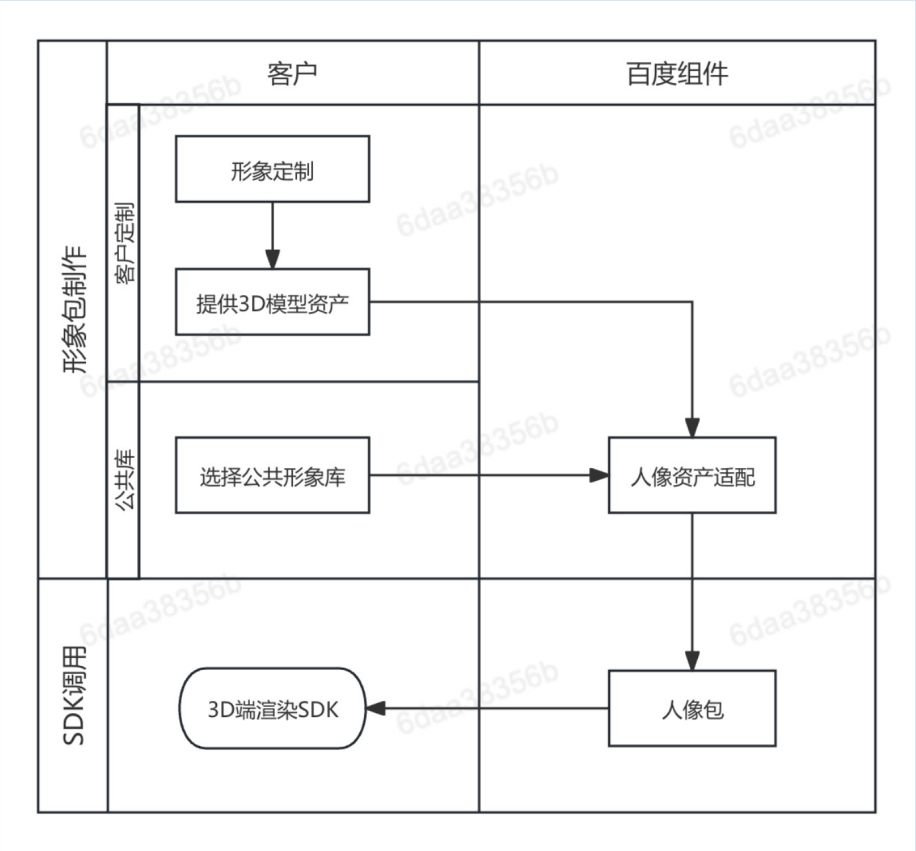
- 1、形象包制作：分为“定制形象”和“公共库形象”两种。

定制形象

用户提供设计原画或已经建模好的3D人物模型。为保证人物模型可以被端渲染SDK驱动使用，百度侧进行人像模型适配，预计需要3-4周时间。人像适配过程中允许用户调整和修改，确保人像符合用户需求。

公共形象库

用户从公共形象库中选择心仪的形象，公共形象参考本文最后《附件.3D端渲染人像库》。为保证人物模型可以被端渲染SDK驱动使用，百度侧进行人像模型适配，预计需要3-4周时间。



形象定制及适配费用相关问题，可咨询：[咨询单](#)



SDK内部模块和流程

内部模块

- AvatarEngine: 人像模块，负责人像的显示，动作口型
- RenderEngine：渲染模块，用于编排指令，播放音频及调用人像

模块流程：

```
AvatarEngine->RenderEngine
```

初始化

需要输入组件平台的token，token生成方式，[接口通用说明](#)

鉴权 token 生成方法参考上述文档或者Demo中的TokenUtil.generateToken方法

为了避免密钥泄露，建议由开发者的后端服务保存 AppId/AppKey 生成 token，然后通过客户端和服务端间的可信链路传输

token

SDK 集成

SDK只依赖网络权限，用于连接tts和口型驱动服务

AndroidManifest.xml文件

```
<uses-permission android:name="android.permission.INTERNET" />
<uses-permission android:name="android.permission.ACCESS_NETWORK_STATE" />
```

依赖库

本地库 app\avatar-libs下

1. arplay-release-xxx.aar

2. avatar.jar

开源库：'com.google.code.gson:gson:2.11.0'

备注：目前高版本的gradle中集成本地aar文件的方式主要有两种

1. 如demo，在外层的settings.gradle中添加flatDir dirs("app/avatar-libs")
2. 新建module，build.gradle改为configurations.maybeCreate("default") 及 artifacts.add("default", file('arplay-release-xx.aar'))

推荐UI集成方式

参考demo中的DigitalHumanDisplayFragment，使用TextureView进行渲染

```
private void initDigitalHumanDisplayUI() {
    digitalHumanDisplayFragment = new DigitalHumanDisplayFragment();
    FragmentManager fragmentManager = getSupportFragmentManager();
    FragmentTransaction fragmentTransaction = fragmentManager.beginTransaction();
    fragmentTransaction.add(R.id.fcv, digitalHumanDisplayFragment).commit();
}
```

SDK具体参数及回调

调用流程

- DigitalHumanSettings设置参数，包含上文的license文件的设置
- OpenEngineCallback 设置上面2个模块的回调
- DigitalHumanEngine 使用DigitalHumanSettings和EngineCallback 初始化，初始调用init及onSurfaceTextureAvailable方法
- 等待onModelLoaded回调
- DigitalHumanEngine 控制数字人行为

初始化DigitalHumanEngine

DigitalHumanEngine.init方法


```

/**
 * 初始化数字人
 *
 * @param context
 * @param settings 数字人配置
 * @param engineCallback 多个callback合一
 * @throws ConfigInputException 配置错误
 */
public void init(@NonNull Context context, @NonNull DigitalHumanSettings settings,
                EngineCallback engineCallback) throws ConfigInputException

```

DigitalHumanSettings设置参数：

```

private DigitalHumanSettings createDigitalHumanSettings(String token, String urlPrefix) {
    DigitalHumanSettings digitalHumanSettings = getDigitalHumanSettings();
    try {
        digitalHumanSettings.setUrlPrefix(urlPrefix);
        PlatConfig platConfig = digitalHumanSettings.getPlatConfig();
        platConfig.setToken(token);
        Avatar3DConfig avatar3DConfig = digitalHumanSettings.getAvatarConfig();
        // asset 目录下的xuniren-avatar-debug-0004.license 鉴权文件
        avatar3DConfig.setLicenseFilename("xuniren-avatar-debug-0004");
        // 渲染等级, HIGH使用高精人像, MIDDLE和LOW使用普通, 部分人像无渲染等级区别
        avatar3DConfig.setDisplayQualityLevel(DisplayQualityLevel.HIGH);
        // 模型文件路径, demo中从assets目录复制到该目录
        avatar3DConfig.setModelPath(caseCacheDir.getAbsolutePath());
        // 设置最大并发TTS数量, 注意组件服务端有最大tts并发设置, 所有进行tts发音的设备的实时并发数不能超过这个
        // 并发, 一般设置1或者2
        digitalHumanSettings.getRenderEngineConfig().setMaxTtsConcurrency(MAX_TTS_CONCURRENCY);
    } catch (InputException e) {
        Log.e(TAG, "Fail to init digitalHumanSettings ", e);
    }
    return digitalHumanSettings;
}

```

EngineCallback 初始化：

```

// SplashActivity.initDigitalHumanEngine

EngineCallback engineCallback = new OpenEngineCallback(new DemoAvatarCallback(), new
DemoRenderEngineCallback(getApplicationContext()));

```

DigitalHumanEngine 使用DigitalHumanSettings和EngineCallback 初始化：

```

// SplashActivity.initDigitalHumanEngine
// 检查
DigitalHumanEngine digitalHumanEngine = DigitalHumanEngine.getInstance();
digitalHumanEngine.checkDigitalHumanSettings(digitalHumanSettings, engineCallback);

// 初始化
digitalHumanEngine.init(this, digitalHumanSettings, engineCallback);

//init 调用后需要调用下文的DigitalHumanEngine.onSurfaceTextureAvailable, 完成初始化

```

初始化Texture和生命周期

输入的TextureView的宽高比必须按照9:16，否则渲染画布的宽度会按照高度同比计算
参考Demo中的DigitalHumanDisplayFragment，layout中布局固定了9:16的宽度

- DigitalHumanEngine.onSurfaceTextureAvailable方法，用于监听SurfaceView初始化
- DigitalHumanEngine.onPause方法，绑定activity的onPause事件
- DigitalHumanEngine.onResume方法，绑定activity的onResume事件
- DigitalHumanEngine.destroy方法，销毁

```
/**
 * TextureView初始化成功
 * TextureView.SurfaceTextureListener 透传
 *
 * @param surface
 * @param width
 * @param height
 * @param requestOrientation
 */
public void onSurfaceTextureAvailable(@NonNull SurfaceTexture surface, int width, int height, int requestOrientation)

/**
 * android onPause事件透传
 */
public void onPause()

/**
 * android onResume事件透传
 */
public void onResume()

/**
 * 销毁SDK
 */
public void destroy() {
```

示例

1. 监听onSurfaceTextureAvailable，调用digitalHumanEngine.onSurfaceTextureAvailable

```
private class DisplaySurfaceTextureListener implements TextureView.SurfaceTextureListener {

    @Override
    public void onSurfaceTextureAvailable(@NonNull SurfaceTexture surface, int width, int height) {
        if (getActivity() == null) {
            throw new DemolInvalidProcessException("activity is not attached");
        }
        lastTextureSurface = surface;
        lastTextureSize = new Size(width, height);
        digitalHumanEngine.onSurfaceTextureAvailable(surface, width, height, getActivity().getRequestedOrientation());
    }
}
```

2. textureView 注册事件，注册TextureView setOnTouchListener

```
private void initView(ViewGroup rootView) {  
    TextureView textureView = rootView.findViewById(R.id.texture_view);  
    textureView.setOpaque(false);  
    textureView.setSurfaceTextureListener(new DisplaySurfaceTextureListener());  
    textureView.setRotation(0);  
}
```

3. 生命周期相关调用

```
@Override  
public void onDestroy() {  
    super.onDestroy();  
    digitalHumanEngine.destroy();  
}  
  
@Override  
public void onPause() {  
    super.onPause();  
    digitalHumanEngine.onPause();  
}  
  
@Override  
public void onResume() {  
    super.onResume();  
    digitalHumanEngine.onResume();  
}
```

DigitalHumanEngine 数字人控制方法

参考demo中的AvatarControllerFragment

- DigitalHumanEngine.mute 方法，数字人静音
- DigitalHumanEngine.unMute方法，数字人取消静音

```
/**  
 * 静音  
 */  
public void mute()  
  
/**  
 * 取消静音  
 */  
public void unMute()
```

DigitalHumanEngine 数字人播放方法

参考demo中的InteractiveControllerFragment

- DigitalHumanEngine.playAnim方法

注意，目前数字人的动作可以随时触发，随时打断当前动作。动作流程与播放无关。

1. 可以在字幕收到时，自行触发动作
2. 智能动作开启时，会自带动作，注意此时不要发动作，以免打断

```
/**
 * 播放动画
 * @param animName
 * @param requestId
 */
public void playAnim(String animName, String requestId) ;
```

- TtsConfig类，用于下文的render里的tts设置和智能动作

```
public class TtsConfig {

    /**
     * 发音人
     */
    @With
    private String person = "5116";

    /**
     * 语速，取值范围：0-15
     */
    private int speed = 5;

    /**
     * 音调，取值范围：0-15
     */
    private int pitch = 5;

    /**
     * 组件平台，是否开启智能动作
     */
    @JsonIgnore
    private boolean autoAnimoji = false;
}
```

- DigitalHumanEngine.render方法，播报文本或者上传音频
- DigitalHumanEngine.interrupt方法，打断播报&动作

```

/**
 * 播放文本，使用默认的TtsConfig
 *
 * @param text
 * @param requestId
 */
public void render(String text, String requestId);

/**
 * 播放文本
 * 当使用LITE版克隆音色及公共音色库中支持多语言的音色时，支持自动识别输入语言来进行播报。例：当输入俄文
 * 时，使用俄语播报。
 * @param text
 * @param requestId
 * @param ttsConfig 见TtsConfig类
 */
public void render(String text, String requestId, TtsConfig ttsConfig);

/**
 * 播放音频流
 * @param audioInputStream
 * @param requestId
 */
public void render(IAudioStream audioInputStream, String requestId);

/**
 * 打断当前播放&动作
 */
public void interrupt();

```

IAudioStream 用于SDK回调获取InputStream，注意SDK内部有重试机制，

```

public interface IAudioStream extends AutoCloseable {

    /**
     * 如果请求过程中失败的话 这里可能会多次调用，每次需要一个从头开始的InputStream
     * @return
     */
    InputStream getReadableInputStream();

    /**
     * 需要关闭当前的InputStream
     * @throws Exception
     */
    @Override
    void close() throws Exception;

    /**
     * 当前流读完
     * @param isAllRead
     */
    void onReadFinish(boolean isAllRead);
}

```

 DigitalHumanSettings 总设置参数

EngineCallback 回调

EngineCallback 内实际是2个回调，均为IOnErrorCallback或IOnWarningCallback

- IOnErrorCallback.onError 人像初始化前的流程， 有此回调后SDK流程终止
- IOnWarningCallback.onWarning 不影响流程的警告， SDK初始化成功后，网络相关的问题以onWarning输出

通用报错回调：

```
public interface IOnErrorCallback extends IOnWarningCallback {

    /**
     * 人像初始化前的流程， 有此回调后SDK流程终止
     *
     * @param name 保存名称或者id
     * @param code SDK错误码
     * @param subCode SDK子错误码，通常为依赖模块或者接口的错误码
     * @param message 具体报错信息
     * @param exceptionClassName 异常类名
     */
    void onError(String name, int code, int subCode, String message, String exceptionClassName);
}

public interface IOnWarningCallback {

    /**
     * 不影响流程的警告
     * SDK初始化成功后，网络相关的问题以onWarning输出
     *
     * @param name 名称或者id
     * @param code 错误码
     * @param subCode 子错误码，通常为依赖模块或者接口的错误码
     * @param message 具体报错信息
     * @param exceptionClassName 异常类名
     */
    void onWarning(String name, int code, int subCode, String message, String exceptionClassName);
}
```

处理背景图片

SDK渲染的TextureView为透明， 可以参考demo自行使用ImageView作为背景

AvatarEngine模块

AvatarEngine负责人像的显示，动作口型，获取AvatarConfig对象：

digitalHumanSettings.getAvatarConfig()

方法	作用	示例/默认	备注
displayQualityLevel	渲染等级	DisplayQualityLevel.HIGH	其它选项DisplayQualityLevel.MIDDLE DisplayQualityLevel.LOW
licenseFilename	鉴权文件名。必须保存在assets目录下	xuniren-avatar-debug-0004	不需要.license后缀

IAvatarCallback 回调

```
public interface IAvatarCallback extends IOnErrorCallback {

    /**
     * 模型加载开始
     *
     * @param figureName 人像名称
     */
    void onStart(String figureName);

    /**
     * 模型开始加载
     */
    void onModelLoaded();

    /**
     * 应用服务端的人像配置成功
     *
     * @param characterConfig 服务端的人像配置
     */
    void onFigureConfigToApply(CharacterConfig characterConfig);

}
```

RenderEngine模块

RenderEngine负责播放，任务的编排，从而调用AvatarEngine，获取RenderEngineConfig对象：

```
digitalHumanSettings.getRenderEngineConfig()
```

方法	作用	示例/默认	备注
setRenderWebsocketUrl	渲染服务地址	ws://10.158.95.141/v1/arkit/render	
arrangeWorkerConfig.renderTextQueueSize	播放的语句数量的队列。调用一次render() 或者大模型返回一句话算一次	100	
arrangeWorkerConfig.arrangeQueryTaskConfig.maxTaskRetryTimes	query重试的次数	3	

IRenderEngineCallback 回调

调用时，如render中填入自定义的requestId，可以与回调中的requestId对应，接口返回的requestId会以传入的requestId开头。

流程：

下载tts数据->tts相关回调

播放->播放相关回调

连接下载相关

```

/**
 * /v1/arkit/render websocket打开
 *
 * @param requestId 同render和query请求的requestId
 */
void onTtsWebsocketConnect(String requestId);

/**
 * /v1/arkit/render websocket关闭
 *
 * @param requestId 同render和query请求的requestId
 * @param code websocket关闭码
 * @param reason websocket 关闭原因
 * @param isNormal websocket是否正常关闭
 */
void onTtsWebsocketClose(String requestId, int code, String reason, boolean isNormal);

```

播放相关

```

/**
 * /v1/arkit/render 收到字幕
 *
 * @param requestId 同render和query请求的requestId
 * @param subtitleBody 字幕内容
 */
void onTtsSubtitle(String requestId, ArSubtitleBody subtitleBody);

/**
 * 渲染开始
 *
 * @param requestId
 * @param isLocalAnimTask 如果是本地动作，则为true
 */
void onRenderStart(String requestId, boolean isLocalAnimTask);

/**
 * 渲染结束
 *
 * @param requestId
 * @param isLocalAnimTask
 */
void onRenderComplete(String requestId, boolean isLocalAnimTask);

/**
 * 动作发起后，会收到start回调，如果这个动作被打断，则不会收到end回调
 *
 * @param requestId
 * @param animName
 */
void onAnimStart(String requestId, String animName);

/**
 * 动作发起后，会收到start回调，如果这个动作被打断，则不会收到end回调
 *
 * @param requestId
 * @param animName
 */
void onAnimEnd(String requestId, String animName);

```


🔗 缓存

如果遇见高频播放文本，可以缓存，一个是tts下载后的数据可以缓存，一个是播放后的数据可以缓存，推荐播放后的。

demo中有利用返回的SimpleRenderElement 流，进行缓存的示例，是否缓存及缓存的逻辑由调用方控制 注意字幕也在缓存中，因此要注意字幕切分的选项设置

demo中有使用onTtsAudioPlayed进行缓存的示例，相关代码见DemoRenderEngineCallback和TtsCacheManager

```
/**
 * /v1/arkit/render websocket收到数据
 *
 * @param requestId 同render和query请求的requestId
 * @param element 渲染元素
 */
void onTtsDataDownloaded(String requestId, SimpleRenderElement element);

/**
 * 音频或者动作被播放
 *
 * @param requestId 同render和query请求的requestId
 * @param element 渲染元素
 */
void onTtsAudioPlayed(String requestId, SimpleRenderElement element);
```

🔗 错误码

```
public interface IExceptionCode {

    /**
     * 未知业务相关错误
     */
    int ERROR_EXCEPTION_NOT_BUSINESS = 2;

    /**
     * 业务RUNTIME报错
     */
    int ERROR_EXCEPTION_BUSINESS_RUNTIME = 3;

    /**
     * 内部报错
     */
    int ERROR_EXCEPTION_BUSINESS_INTERNAL = 9;

    /**
     * 内部报错
     */
    int ERROR_EXCEPTION_BUSINESS_UNKNOWN = 10;

    /**
     * DigitalHumanSetting.setUrlPrefix url schema 错误
     */
    int ERROR_INPUT_URL_PREFIX_SCHEMA = 411;

    /**
     * 发送render和query时，引擎人像未准备完成
     */
    int ERROR_ENGINE_AVATAR_NOT_PREPARED = 801;
```

```
/**
 * 调用了改平台没有的功能，请勿调用
 */
int ERROR_ENGINE_PLATFORM_NOT_ALLOWED = 802;

/**
 * 内部解密错误
 */
int ERROR_CODEC_DECRYPT_FAIL = 901;

/**
 * platConfig.appId为空
 */
int ERROR_SETTINGS_PLAT_APPID_NULL = 1011;

/**
 * platConfig.appKey为空
 */
int ERROR_SETTINGS_PLAT_APPKEY_NULL = 1012;

/**
 * platConfig.generateUrl为空
 */
int ERROR_SETTINGS_PLAT_GENERATEURL_NULL = 1013;

/**
 * platConfig.appInfoUrl为空
 */
int ERROR_SETTINGS_PLAT_APPINFOURL_NULL = 1014;

/**
 * 填写的PlatConfig鉴权为空
 */
int ERROR_SETTINGS_PLAT_TOKEN_NULL = 1015;

/**
 * [组件平台] 模型文件路径
 */
int ERROR_SETTINGS_AVATAR_MODEL_PATH_NOT_CORRECT = 1016;

/**
 * 渲染用的websocket链接为空， 可以调用digitalHumanSettings.setUrlPrefix
 */
int ERROR_SETTING_RENDERENGINE_WEBSOCKET_URL_NULL = 1017;

/**
 * callback为空
 */
int ERROR_SETTING_CALLBACK_NULL = 1021;

/**
 * platCallback为空
 */
int ERROR_SETTING_PLAT_CALLBACK_NULL = 1022;

/**
 * downloadCallback为空
 */
int ERROR_SETTING_DOWNLOAD_CALLBACK_NULL = 1023;

/**
 * avatarCallback为空
```

```
*/
int ERROR_SETTING_AVATAR_CALLBACK_NULL = 1024;

/**
 * renderEngineCallback为空
 */
int ERROR_SETTING_RENDERENGINE_CALLBACK_NULL = 1025;

/**
 * 模型文件目录不可写
 */
int ERROR_SETTINGS_AVATAR_MODEL_PATH_NOT_WRITEABLE = 1031;

/**
 * http请求IO异常
 */
int ERROR_HTTP_IO_EXCEPTION = 2001;

/**
 * http返回值不为200
 */
int ERROR_HTTP_RESPONSE_CODE_NOT_200 = 2002;

/**
 * http返回值为200，但是内容是空
 */
int ERROR_HTTP_RESPONSE_EMPTY = 2003;

/**
 * 发送http请求前，序列化json失败
 */
int ERROR_HTTP_SERIALIZE_JSON = 2010;

/**
 * 发送http请求后，反序列化json失败
 */
int ERROR_HTTP_DESERIALIZE_JSON = 2011;

/**
 * arinfo接口，服务端返回配置内容为空
 */
int ERROR_PLAT_HTTP_APPINFO_EMPTY = 2101;

/**
 * platConfig.getAppInfoUrl() 里设置url不合法
 */
int ERROR_PLAT_HTTP_URL_INCORRECT = 2102;

/**
 * 修改服务端人像，不能重新加载配饰
 */
int ERROR_PLAT_FIGURE_NOT_SAME = 2103;

/**
 * 内部错误，arinfo配置不能被序列号
 */
int ERROR_PLAT_DEEPCOPY_ARINFO = 2104;

/**
 * 内部错误，platengine状态未准备好
 */
int ERROR_PLAT_AVATAR_NOT_PREPARED = 2105;
```

```
/**
 * 服务端返回的arinfo接口，successful为false
 */
int ERROR_PLAT_RESPONSE_NOT_SUCCESSFUL = 2111;

/**
 * 服务端generate接口，返回的token为空
 */
int ERROR_PLAT_GENERATE_TOKENENC_EMPTY = 2112;

/**
 * 服务端generate接口，返回的token为空
 */
int ERROR_PLAT_TOKEN_NULL = 2113;

/**
 * 服务端generate接口，返回的加密token为空
 */
int ERROR_PLAT_GENERATE_TOKEN_DECRYPTY_ENC_EMPTY = 2114;

/**
 * 下载目录不可写
 */
int ERROR_DOWNLOAD_DIRECTORY_NOT_WRITABLE = 2201;

/**
 * 内部错误
 */
int ERROR_DOWNLOAD_URL_CALLBACK_EMPTY = 2203;

/**
 * 下载文件的md5校验失败
 */
int ERROR_DOWNLOAD_MD5_CHECK = 2207;

/**
 * 配置的ar目录不存在，请检查文件目录
 */
int ERROR_AVATAR_CASE_DIR_NOT_EXIST = 2302;

/**
 * 配置的ar目录下，不存在../avatar/{人像id}/config.lua
 */
int ERROR_AVATAR_FILE_CONFIG_LUA_NOT_EXIST = 2303;

/**
 * 人像初始化鉴权失败
 */
int ERROR_AVATAR_AUTH_EXCEPTION = 2311;

/**
 * 人像驱动lua脚本返回不为0的值
 */
int ERROR_AVATAR_LUA_MESSAGE_CALLBACK = 2312;

/**
 * 人像引擎返回错误
 */
int ERROR_AVATAR_DU_MIX_CALLBACK = 2313;

/**
```

```

    * 人像引擎onEngineSetup异常
    */
    int ERROR_AVATAR_DU_MIX_CALLBACK_SETUP = 2314;

/**
    * Avatar3DEngine 不是可以播放的状态
    */
    int ERROR_AVATAR_STATUS_NOT_NORMAL = 2315;

/**
    * 人像资源解析错误
    */
    int ERROR_AVATAR_RESOURCE_PARSE_PART_JSON = 2321;

/**
    * 人像资源不存在
    */
    int ERROR_AVATAR_RESOURCE_VALUE_NOT_EXIST = 2322;

/**
    * model_android.json 文件不存在
    */
    int ERROR_AVATAR_RESOURCE_ANIM_CONFIG_NOT_EXIST = 2331;

/**
    * model_android.json 文件存在但不可读
    */
    int ERROR_AVATAR_RESOURCE_ANIM_CONFIG_NOT_READABLE = 2332;

/**
    * 人像的配饰的模型目录不存在
    */
    int ERROR_AVATAR_RESOURCE_PART_DIR_NOT_READABLE = 2333;

/**
    * [组件平台]模型文件model目录不存在
    */
    int ERROR_SIMPLE_CHARACTER_CONFIG_MODEL_PATH_NOT_EXIST = 2341;

/**
    * 模型文件model目录下计算人像错误
    */
    int ERROR_SIMPLE_CHARACTER_CONFIG_AUTO_FIND_FIGURE_NAME_FAILED = 2342;

/**
    * 打开用户提供的音频流报错
    */
    int ERROR_ARRANGE_RENDER_AUDIO_STREAM_OPEN_FAILED = 2401;

/**
    * websocket连接超时，默认5s
    */
    int ERROR_WEBSOCKET_CONNECT_TIMEOUT = 2411;

/**
    * interrupt 后，原有线程消耗
    */
    int ERROR_ARRANGE_RENDER_WEBSOCKET_INTERRUPTED = 2412;

/**

```

```
    * 渲染队列过长
    */
    int ERROR_ARRANGE_WORKER_QUEUE_FULL = 2501;

    /**
     * 渲染websocket重试过多次数
     */
    int WARNING_ARRANGE_WORKER_WEBSOCKET_RETRY_TOO_MANY_TIMES = 2502;

    /**
     * 渲染一帧延时大于40ms
     */
    int WARNING_FRAME_DELAY_GREATER_40 = 2511;

    /**
     * 一个render任务，渲染延时综合大于120ms
     */
    int WARNING_CUMULATE_FRAME_DELAY_TOO_LARGE = 2512;

    /**
     * 渲染接口，鉴权失败
     */
    int ERROR_RENDER_AUTHENTICATION_FAILED = 2601;

    /**
     * 渲染接口，尝试多次失败
     */
    int ERROR_RENDER_CONNECT_TOO_MANY_TIMES = 2602;

    /**
     * babel 的websocket链接端开，可能是网络原因导致
     */
    int ERROR_ARRANGE_BABEL_QUERY_WEBSOCKET_NOT_OPENED = 2701;

    /**
     * babel提问失败
     */
    int ERROR_ARRANGE_BABEL_QUERY_FAILED = 2702;

}
```

3D数字人交互SDK-iOS版

简介

本文说明SDK的集成使用以及接口调用方式

整体架构&功能简介

- 客户端或页面集成SDK加载形象包，本地驱动渲染数字人。
- 支持音频和文本驱动:
- 音频可以使用真人音频或自行接入的第三方的TTS音频。
- 文本驱动SDK内部会调用数字人开放平台TTS服务驱动数字人，
- 文本驱动可以开启智能动作开关，实现数字人根据播报语义智能触发动作。
- 3D数字人形象包可切换机位，动作驱动，切换配饰服装（如有多套配置服装）



下载地址

版本号	下载地址	更新时间	SDK包体大小
2.0.0.4	下载地址	2025.01.13	6mb（不包含人像包大小）

🔗 形象定制（使用公共形象可以跳过）

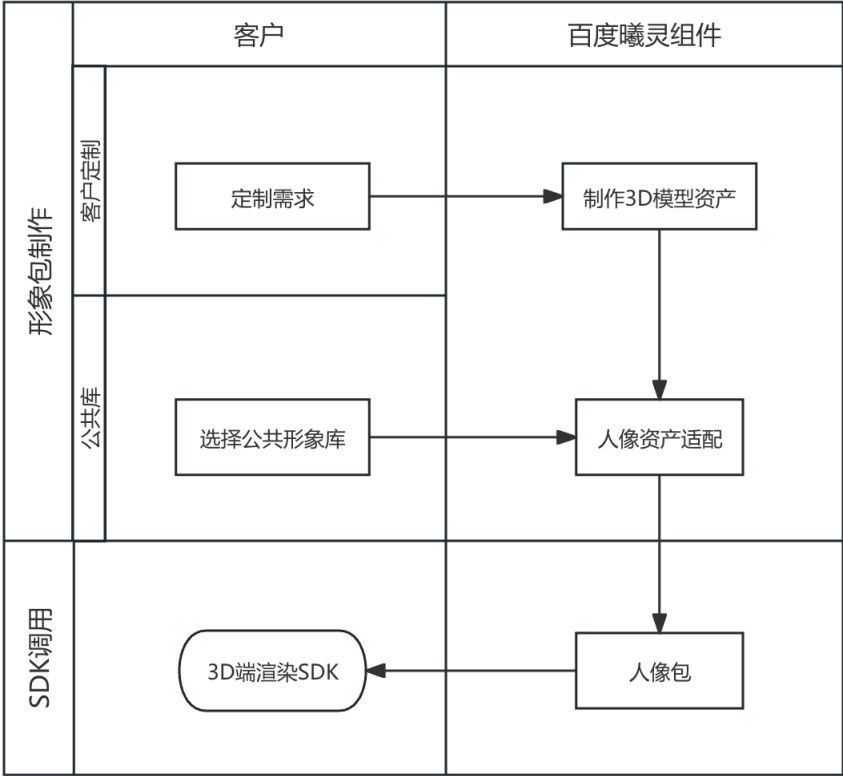
1、形象包制作：分为“定制形象”和“公共库形象”两种。

• 定制形象

用户提供设计原画或已经建模好的3D人物模型。为保证人物模型可以被端渲染SDK驱动使用，百度侧进行人像模型适配，预计需要3-4周时间。人像适配过程中允许用户调整和修改，确保人像符合用户需求。

• 公共形象库

用户从公共形象库中选择心仪的形象，公共形象参考本文最后《附件·公共人像库》。

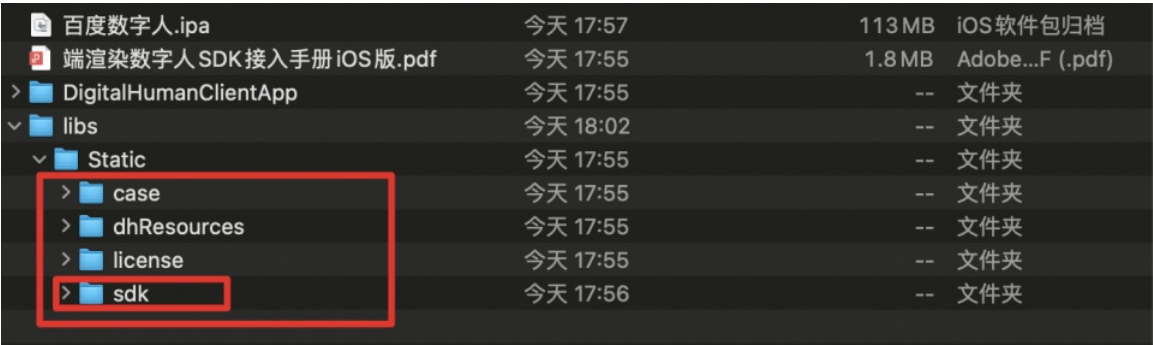


形象定制及适配费用相关问题，可咨询：[咨询单](#)



一、SDK接入和集成

1.1、将下载得到的framework文件和a文件在framework，librarties，and embedded content中进行导入

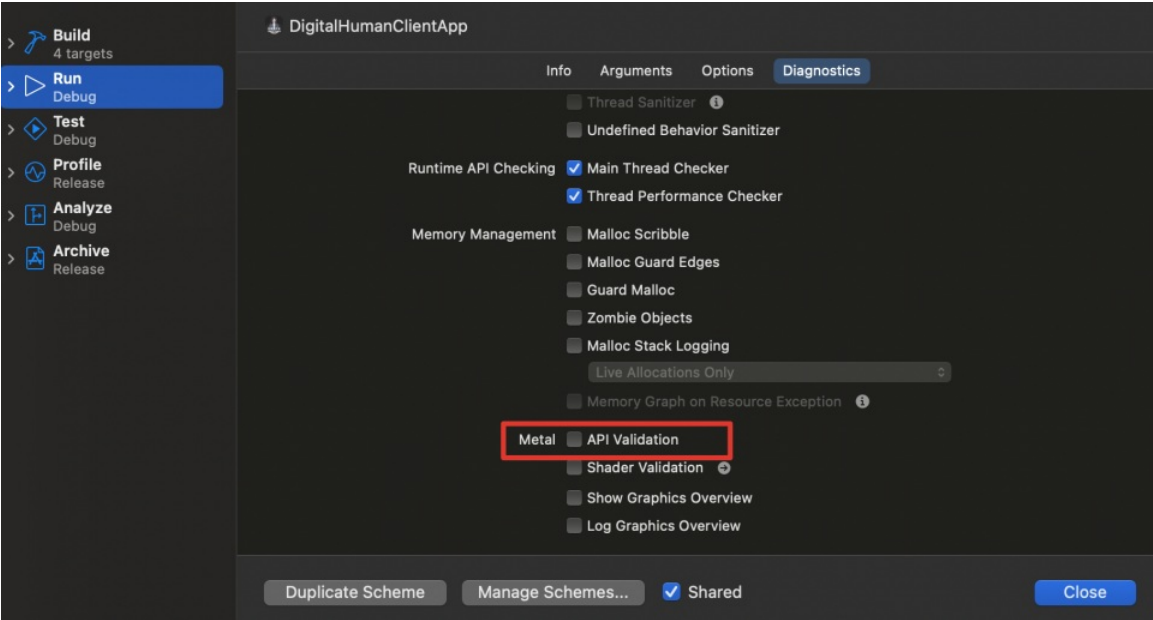


1.2、添加以下系统库

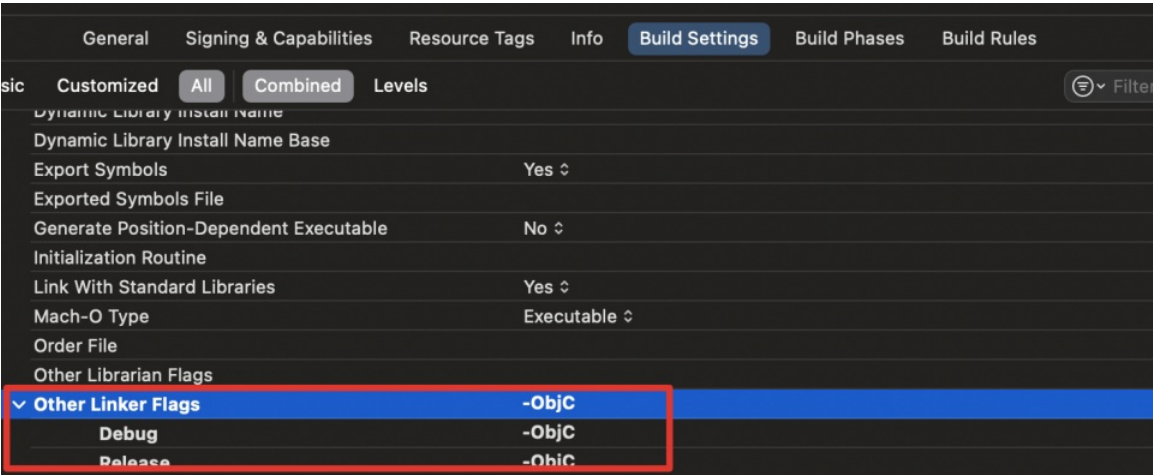
Framework	描述
libc++.tbd	提供对C/C++特性支持
GLKit.framework	GL支持
OpenGLES.framework	OpenGL 支持
libcucore.tbd	
libz.tbd	提供gzip支持
libnetwork.tbd	提供网络支持

1.3、在工程中配置如下参数

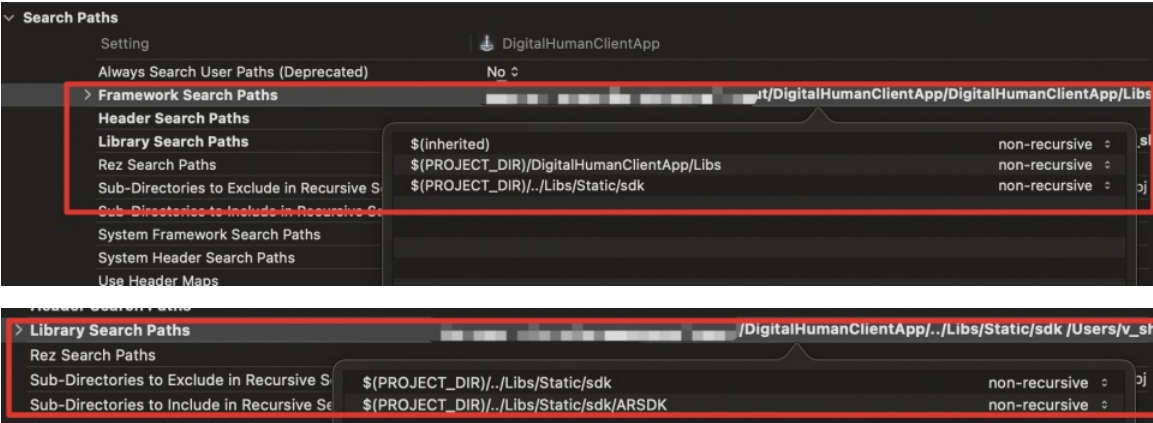
1.3.1、将 API Validation 取消选中



1.3.2、将other Linker Flages 设置为-ObjC



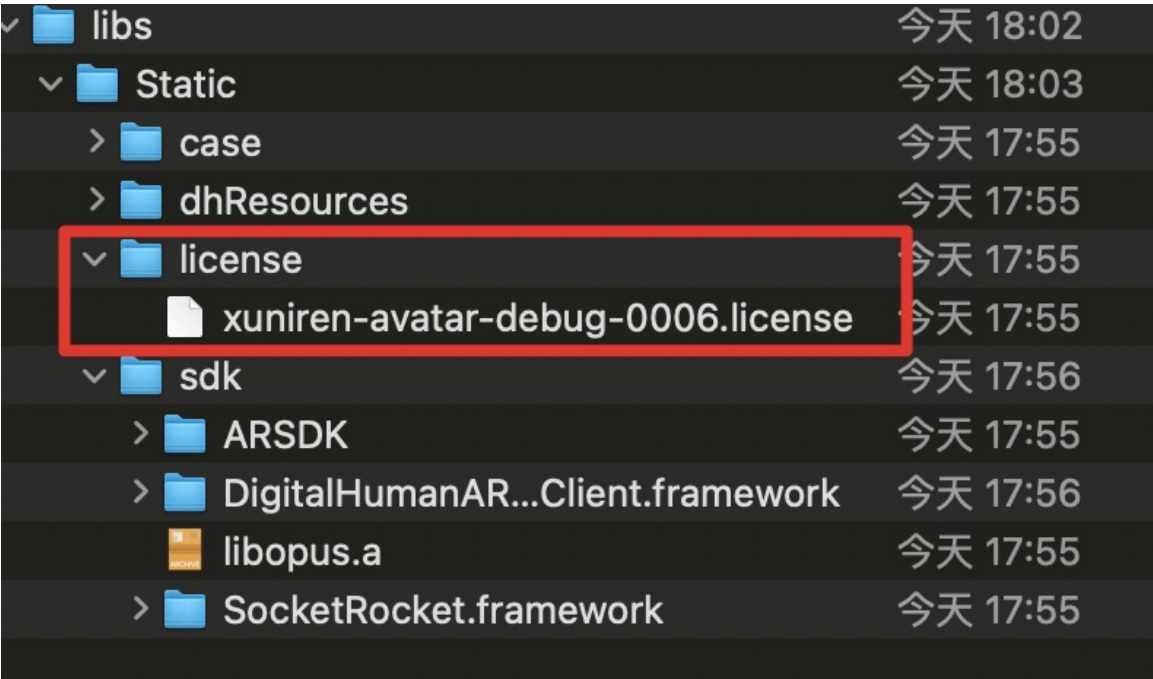
1.3.3、库引用地址



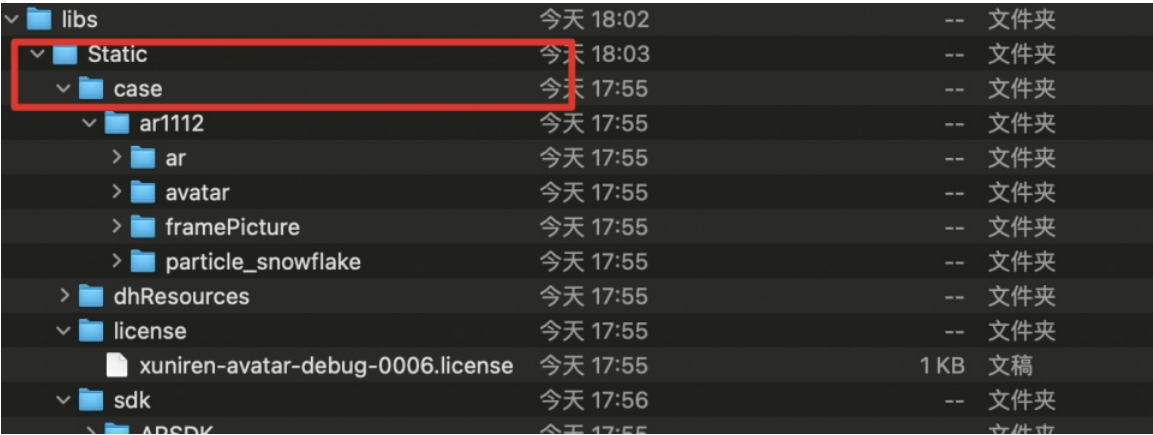
1.4、资源文件导入

1.4.1、license 文件资源导入（必须作为资源导入， 否则授权则不成功， 如果有新的授权文件则更新到license 文件夹下）

license需要额外申请， 提供Bundle Identifier 给到百度方， 用于生成license文件



1.4.2、人像资源导入 (可自定义地址， 初始化时传入自定义地址即可， 地址获取至case目录)



1.5、SDK性能要求

支持机型信息：

- 安卓：支持CPU骁龙730G/骁龙855/联发科天玑900/天玑820 同等性能及以上芯片
- iOS：支持iPhone8 系统版本13.3以及以上。

二、SDK 调用以及API 说明

2.1、初始化

```
-(instancetype)initWithFrame:(CGRect)frame andARVSSettings:(DigitalHumanARVSSettings *)settings
```

参数	含义	是否必填	备注
frame	数字人view显示坐标	是	
settings	加载ar数字人、babel、render 的配置信息	是	参数参考DigitalHumanARVSSettings 模块

DigitalHumanARVSSettings 参数

参数	含义	是否必填	备注
appId	应用的appId	是	咨询百度支持人员
appKey	应用的appKey	是	
realmName	tts合成域名	是	
arParams	ar数字人配置参数	是	参考arParams 参数
renderEngine	render播报配置参数	是	参考renderEngine 参数

arParams 参数

参数	含义	是否必填	备注
modelPath	模型地址	是	
figure	人像名称	是	
camera	机位	是	
deviceType	渲染等级	是	低端机：iPhone_Low 中端机：iPhone_Middle 高端机：iPhone_Heavy
licenseName	授权文件名称	是	咨询百度支持人员
licenseKey	授权key	是	

renderEngine参数

参数	含义	备注
renderToken	token	
characterImage	人像id	用于标识人像，目前只用于开启智能动作功能
autoAnimoji	是否开启智能动作匹配	YES: 开启 NO: 不开启 开启条件： autoAnimoji = YES 且characterImage 为有效值则打开，否则无法生效
ttsParams	tts参数	示例： {@"person":@5116,@"pitch":@5,@"speed":@5}

调用示例

```
_cloudView = [[DigitalHumanARVSVIEW alloc] initWithFrame:self.view.frame andARVSSettings:_dhSettings];
_cloudView.delegate = self;
_cloudView.backgroundColor = [UIColor clearColor]; // 初始化时设置背景为透明
[self.view addSubview:_cloudView];
```

2.2、获取版本信息

```
+ (NSString *)version
```

调用示例

```
[DigitalHumanARVSVIEW version]
```

2.3、数字人暂停

```
- (void)pause
```

调用示例

```
[_cloudView pause];
```

2.4、数字人继续

```
- (void)resume
```

调用示例

```
[_cloudView resume];
```

2.5、render播放相关接口

2.5.1、文本播报

```
- (void)sendTextRender:(NSString *)message
```

参数	含义	是否必填	备注
message	播报文本	是	

调用示例

```
[_cloudView sendTextRender:text];
```

2.5.2、音频文件播报

```
- (void)sendAudioStream:(NSString *)input
```

参数	含义	是否必填	备注
input	音频文件路径	是	当使用LITE版克隆音色及公共音色库中支持多语言的音色时，支持自动识别输入语言来进行播报。例：当输入俄文时，使用俄语播报。

调用示例

```
NSString *micPath = [[NSBundle mainBundle] pathForResource:@"asr_0501-iPhone13Promax-ljjiaqi_kongtiao_mic_01"
ofType:@"pcm" inDirectory:@"dhResources"];
[_cloudView sendAudioStream:micPath];
```

2.6、打断

```
- (void)interrupt
```

调用示例

```
[_cloudView interrupt];
```

2.7、播放动作

```
- (void) playAnim:(NSString *) animName
```

参数	含义	是否必填	备注
animName	播报动作ID	是	

调用示例

```
[_cloudView playAnim:_animationList[_animationIndex]];
```

2.8、设置idle

```
- (void) setArdle
```

调用示例

```
[_cloudView setArdle];
```

2.9、更新render参数

```
-(void) updateRenderParams:(RenderEngine *) renderEngine;
```

参数	含义	是否必填	备注
renderEngine	合成相关参数	是	

renderEngine

参数	含义	备注
renderToken	token	不更新token 项无需调用
ttsParams	tts参数	不更新tts 项无需调用，调用示例：{"person":@5116,"pitch":@5,"speed":@5}

调用示例

```
RenderEngine * renderEngine = self.dhSettings.renderEngine;
// 更新token
renderEngine.renderToken = [DHLanAES generateAuthorizationTokenWithAppld:self.dhSettings.appld
appKey:self.dhSettings.appKey expirationTime:1];
// 更新tts参数
NSDictionary *tts = @{
    @"person":@5116,
    @"pitch":@5,
    @"speed":@5,
};
renderEngine.ttsParams = tts;
[_cloudView updateRenderParams:renderEngine];
```

2.10、切换机位

```
- (void)changeCamera:(NSString*) camParam
```

参数	是否必填	含义
camParam	是	机位ID

调用示例

```
[_cloudView changeCamera:_cameraData[0]];
```

2.11、销毁

```
- (void)destroy
```

调用示例

```
[_cloudView destroy];
```

2.12、DigitalHumanARVSDelegate

```
/**
数字人sdk状态变化回调

@param event 状态名称
@param data 根据状态的不同，绑定的其他数据
*/
@required
- (void)onDigitalHumanARVSEventUpdate:(DigitalHumanARVSEvent)event obj:(id __nullable)data;
```

参数	含义	备注
onDigitalHumanARVSEventUpdate	sdk状态变化回调	参数参考DigitalHumanARVSEvent 模块

DigitalHumanARVSEvent 参数

参数	含义	备注
DH_ARVS_LIB_OPEN	初始化完成	
DH_ARVS_LIB_ERROR	数字人执行错误	
DH_ARVS_LIB_LICENSE	授权状态	数字人授权状态
DH_ARVS_LIB_AR_RENDERING	获取数字人	后台返回数字人页面触发
DH_ARVS_LIB_AR_PAUSE	暂停数字人	退出至后台触发
DH_ARVS_LIB_CLOSED	关闭销毁数字人	
DH_ARVS_LIB_AVATAR_ANIMATION	获取可执行动作列表	

2.13、DigitalHumanARVSEnderDelegate

```
/**
播报状态变化回调

@param event 状态名称
@param data 根据状态的不同，绑定的其他数据
*/
@protocol DigitalHumanARVSEnderDelegate <NSObject>

-(void)onDigitalHumanRenderEventUpdate:(QueryEventStatus) event obj:(id __nullable)data;
```

QueryEventStatus 参数

参数	含义	备注
RENDER_START	开始播放	
RENDER_INTERRUPT	播放打断	
RENDER_FINISH	播放结束	
RENDER_ERROR	播放错误	
SUBTITLE	播报内容	音频驱动无播报内容，忽略此字段

附件. 公共人像库

人像	人像包
Yuki	https://meta-human-editor-prd.bj.bcebos.com/open-api/yuki_双端_250117-3.zip

手语数字人交互SDK-Android版

简介

本文档说明Android端手语数字人SDK使用以及调用方式

一、集成

1、准备工作 **1.1、包名和签名md5获取** 包名：每个 Android 应用均有一个唯一的应用 ID。安卓以 Java 包的形式管理应用。

```
defaultConfig {
    applicationId "com.baidu.ar.avatar.sample"
}
```

具体说明见官方文档：<https://developer.android.com/build/configure-app-module?hl=zh-cn>

签名md5

生成签名文件：<https://developer.android.com/studio/publish/app-signing.html>

方法1：

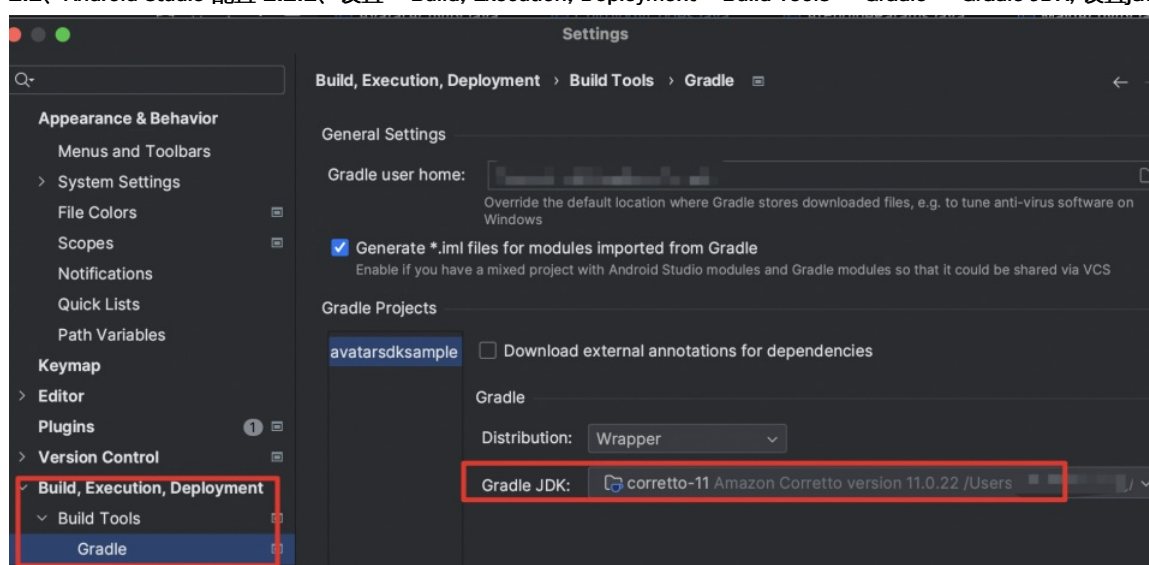
```
命令行输入 keytool -list -v -keystore <签名文件路径>
```

2、SDK 集成 引入SDK aar包： module_core-jnidev-release.aar.aar

2.1、app目录新建libs目录，放入aar包，在build.gradle中增加配置如下

```
implementation 'com.google.code.gson:gson:2.10.1'
implementation files('libs/module_core-jnidev-release.aar.aar')
```

2.2、Android studio 配置 2.2.1、设置-> Build, Execution, Deployment-> Build Tools -> Gradle -> Gradle JDK, 设置java 11



2.3、系统权限配置以及申请 2.3.1、权限配置

```
<uses-permission android:name="android.permission.INTERNET" />
<uses-permission android:name="android.permission.READ_EXTERNAL_STORAGE" />
<uses-permission android:name="android.permission.WRITE_EXTERNAL_STORAGE" />
<uses-permission android:name="android.permission.MODIFY_AUDIO_SETTINGS" />
<uses-permission android:name="android.permission.CHANGE_WIFI_STATE" />
<uses-permission android:name="android.permission.ACCESS_NETWORK_STATE" />
<uses-permission android:name="android.permission.ACCESS_WIFI_STATE" />
<uses-permission android:name="android.permission.READ_EXTERNAL_STORAGE" />
<uses-permission android:name="android.permission.MANAGE_EXTERNAL_STORAGE" />
```

2.3.2、权限申请

参考demo BaseActivity示例，所列举的权限为最低权限要求，如果缺少某一个，会导致sdk出现异常。

二、SDK 调用以及API 说明

1、环境设置

```
void setEnv(Context context, ChirologyCodes.EnvType envType);
```

参数	含义	备注
context	上下文	
envType	Release：正式环境 Develop：测试环境	

调用示例

```
ChirologyAvatar.getInstance().setEnv(MainActivity.this, ChirologyCodes.EnvType.Develop);
```

2、日志开关

```
void setLogType(LogType type)
```

参数	含义	备注
type	NONE：关闭; ERROR：错误信息; WARN：警告信息; DEBUG：debug信息; INFO：info信息; VERBOSE: 全量信息	

调用示例

```
LogTool.setLogType(ChirologyCodes.LogType.NONE);
```

3、SDK 授权相关 3.1、设备授权

appKey，seceretKey为一台设备独有，一个设备要使用手语数字人必须要有一对appKey，seceretKey

```
boolean initLicenseBatch(final Context context, final String appKey,  
                        final String seceretKey, final VirtualAuth.AuthCallBack authCallback)
```

参数	含义	备注
context	上下文	
appKey	设备ak	非空，由百度提供，为一台设备的独有。请向百度申请
seceretKey	设备sk	非空，由百度提供，为一台设备的独有。请向百度申请
authCallback	code：0：成功，其他：失败 subCode：后端返回状态码 msg：提示消息	

调用示例


```
String ak = deviceAppKeyEdit.getText().toString().trim();
String sk = deviceSeceretKeyEdit.getText().toString().trim();
ChirolodyAvatar.getInstance().initLicenseBatch(MainActivity.this, ak, sk, new VirtualAuth.AuthCallBack() {
    @Override
    public void onMessage(int code, int subCode, String msg) {

    }
});
```

3.2、设备授权码激活/续期

当使用新的appKey，seceretKey时必须进行激活，激活方式为使用时长码激活，激活码会随着appKey，seceretKey 提供，使用项目授权方式不需要激活。

```
boolean renewExpirationDate(final Context context, final String activateCode,
    final VirtualAuth.AuthRenewCallBack renewCallBack)
```

参数	含义	备注
context	上下文	
activateCode	激活码	百度提供ak.sk 时携带的激活码
authCallback	code：0：成功，其他：失败 subCode：后端返回状态码 msg：提示消息	

调用示例

```
String trim = activateCodeEdit.getText().toString().trim();
ChirolodyAvatar.getInstance().renewExpirationDate(MainActivity.this, trim, new VirtualAuth.AuthRenewCallBack() {
    @Override
    public void onAuthRenewStatus(int code, int subCode, long expirationTime, String msg) {

    }
});
```

3.3、项目授权

```
boolean initProjectAuth(final Context context, final String projectAk, final String projectSk,
    final VirtualAuth.AuthCallBack authCallback)
```

参数	含义	备注
context	上下文	
projectAk	项目ak	非空，由百度提供，为一个项目独有。请向百度申请
projectSk	项目sk	非空，由百度提供，为一个项目独有。请向百度申请
authCallback	code：0：成功，其他：失败 subCode：后端返回状态码 msg：提示消息	

调用示例

```
String projectAk = projectAkEdit.getText().toString().trim();
String projectsk = projectskEdit.getText().toString().trim();
ChirologyAvatar.getInstance().initProjectAuth(MainActivity.this, projectAk, projectsk, new VirtualAuth.AuthCallBack() {
    @Override
    public void onMessage(int code, int subCode, String msg) {

    }
});
```

3.4、查看设备状态

boolean getOAuthState(final Context context, final VirtualAuth.AuthStateCallBack authStateCallBack)

参数	含义	备注
context	上下文	
authStateCallBack	code : 0 : 成功，其他 : 失败 subCode: 后端返回状态码 activeTime : 生效时间，异常时返回0 expirationTime : 到期时间，异常时返回0 msg : 提示消息	

调用示例

```
ChirologyAvatar.getInstance().getOAuthState(MainActivity.this, new VirtualAuth.AuthStateCallBack() {
    @Override
    public void onAuthStatus(int code, int subCode, long activeTime, long expirationTime, String msg) {

    }
});
```

3.5、换绑设备

当一个授权码还存在有效期，但是设备坏了的情况下，可以使用这个接口进行授权设备的换绑，使用项目授权方式时不可进行换绑。

换绑接口有次数限制且占用总设备数量，非特殊情况不宜调用，最大换绑次数3次

boolean changeDeviceBind(final Context context, final VirtualAuth.ChangeDevCallBack callBack)

参数	含义	备注
context	上下文	
callBack	code : 0 : 成功，其他 : 失败 subCode : 后端返回状态码 msg : 提示消息	

调用示例

```
ChirologyAvatar.getInstance().changeDeviceBind(MainActivity.this, new VirtualAuth.ChangeDevCallBack() {  
    @Override  
    public void onAuthStatus(int code, int subCode, String msg) {  
  
    }  
});
```

4、手语资源 4.1、资源复制

将下载后的chirologyData.zip手动推送到设备的根目录下 (/sdcard)后调用复制资源接口即可

```
void copySignLanguageRes(Context context, CopyResourcesCallback copyResourcesCallback);
```

参数	含义	备注
context	上下文	
copyResourcesCallback	复制回调	

CopyResourcesCallback 参数

参数	含义	备注
onCopySuccess()	复制成功	
onCopyFailed(int code, int subCode, String error)	复制失败	失败原因参考

调用示例

```
ChirologyAvatar.getInstance().copySignLanguageRes(getApplicationContext(), new  
IChirologyAvatar.CopyResourcesCallback() {  
    @Override  
    public void onCopySuccess() {  
  
    }  
  
    @Override  
    public void onCopyFailed(int code, int subCode, String msg) {  
  
    }  
});
```

4.2、资源下载以及复制

```
void downloadCopySignLanguageRes(Context context, String url, CopyResourcesCallback copyResourcesCallback)
```

参数	含义	备注
context	上下文	
url	资源地址	
copyResourcesCallback	复制回调	

CopyResourcesCallback 参数

参数	含义	备注
<code>onCopySuccess()</code>	复制成功	
<code>onCopyFailed(int code, int subCode, String error)</code>	复制失败	失败原因参考

调用示例

```
String url = "资源地址";
ChirologyAvatar.getInstance().downloadCopySignLanguageRes(
    getApplicationContext(), url, new IChirologyAvatar.CopyResourcesCallback() {
        @Override
        public void onCopySuccess() {

        }

        @Override
        public void onCopyFailed(int code, int subCode, String msg) {

        }
    });
```

5、初始化数字人

```
void initAvatar(Context context, ArEngineParams arEngineParams, InitCallBack initCallBack);
```

参数	含义	备注
context	上下文	
arEngineParams	配置参数	
initCallBack	初始化回调	

ArEngineParams 参数

参数	含义	备注
appId	项目的appId	
appKey	项目的appKey	
token	token	
expiredTime	token的有效期	
url	手语翻译地址	
avatarName	人像名称	

InitCallBack 参数

参数	含义	备注
<code>initSuccess()</code>	初始化成功	
<code>initFailed(int code, int subCode, String msg)</code>	初始化失败	失败原因参考

调用示例

```
chirologyAvatar.initAvatar(getApplicationContext(), arEngineParams, new IChirologyAvatar.InitCallBack() {

    @Override
    public void initSuccess() {

    }

    @Override
    public void initFailed(int code, int subCode, String msg) {

    }

});
```

5.1、显示数字人

配置在textureView

```
void setupController(int requestedOrientation, SurfaceTexture surface, int width, int height)
```

参数	含义	备注
requestedOrientation	方位角	
surface	显示的surface	
width	view 宽	
height	view 高	

调用示例

```
public void onSurfaceTextureAvailable(SurfaceTexture surfaceTexture, int width, int height) {
    if (chirologyAvatar != null) {
        chirologyAvatar.setupController(getRequestedOrientation(), surfaceTexture, width, height);
    }
}
```

5.2、计算比例

```
void adjustRenderAspectRatio(float ratio)
```

参数	含义	备注
ratio	比例值	

调用示例

```
// 计算宽高比
float ratio = numerator / denominator;
// 调整渲染的宽高比
if (chirologyAvatar != null) {
    chirologyAvatar.adjustRenderAspectRatio(ratio);
}
```

5.3、更新size

```
void updateOutSize(int width, int height)
```

参数	含义	备注
width	更新的宽	
height	更新的高	

调用示例

```
public void onSurfaceTextureSizeChanged(SurfaceTexture surfaceTexture, int width, int height) {  
    if (chirologyAvatar != null) {  
        chirologyAvatar.updateOutSize(width, height);  
    }  
}
```

6、进入页面

该接口与生命周期强绑定

```
void resume()
```

调用示例

```
protected void onResume() {  
    super.onResume();  
    if (chirologyAvatar != null) {  
        chirologyAvatar.resume();  
    }  
}
```

7、退到后台

该接口与生命周期强绑定

```
void pause();
```

调用示例

```
protected void onPause() {  
    super.onPause();  
    if (chirologyAvatar != null) {  
        chirologyAvatar.pause();  
    }  
}
```

8、DRML 驱动数字人

```
void sendDrmlRender(String drml, float speed, AnimRenderCallBack animRenderCallBack);
```

参数	含义	备注
drml	drml 内容	
speed	播放倍速	默认1.2倍
animRenderCallBack	播放回调	

AnimRenderCallBack 参数

参数	含义	备注
onRenderStart()	开始播报	
onRenderInterrupt()	播报打断	
onRenderFinish()	结束播报	
onRenderError(int code, int subCode, String msg)	播报错误	错误参考

调用示例

```
chirologyAvatar.sendDrmlRender(editTxt.getText().toString().trim(), sped,
                                AvatarActivity.this);
```

9、文本驱动数字人

```
void sendTextRender(String text, float speed, AnimRenderCallBack animRenderCallBack);
```

参数	含义	备注
text	文本内容	
speed	播放倍速	默认1.2倍
animRenderCallBack	播放回调	

AnimRenderCallBack 参数

参数	含义	备注
onRenderStart()	开始播报	
onRenderInterrupt()	播报打断	
onRenderFinish()	结束播报	
onRenderError(int code, int subCode, String msg)	播报错误	错误参考

调用示例

```
chirologyAvatar.sendTextRender(editTxt.getText().toString().trim(), sped,
                                AvatarActivity.this);
```

10、打断

```
void interrupt() ;
```

调用示例

```
chirologyAvatar.interrupt();
```

11、释放

```
void release() ;
```

调用示例

```
chirologyAvatar.release();
```

12、其他

12.1、授权相关错误说明

code	subCode	含义
通用	0	成功
	-1006	授权失败
	-1007	引擎状态异常
查看授权状态	-1001	获取状态失败，SDK状态异常
	-1002	获取状态失败，服务端返回为空
	-1003	获取状态失败，服务端状态不成功
	-1004	获取状态失败，没有返回license
	-1005	获取状态失败，一般是网络异常
激活/续期	-1008	续期失败，发生异常
	-1009	续期失败，SDK状态异常
	-1010	续期失败，后台返回为空
	-1011	续期失败，后台状态不成功
设备换绑	-1012	设备切换失败，SDK状态异常
	-1013	设备切换失败，服务端返回为空
	-1014	设备切换失败，服务端状态不成功
	-1015	设备切换失败，一般是网络异常
项目授权	-1016	获取项目授权信息失败，授权信息解析为空
	-1017	获取项目授权信息失败，服务端状态不成功
	-1018	获取项目授权信息失败，没有返回license
	-1023	ProjectAK为空
	-1024	ProjectSK为空
	-1025	ProjectAK长度错误
	-1026	ProjectSK长度错误
设备授权	-1019	AppKey为空
	-1020	SecretKey为空
	-1021	AppKey长度错误
	-1022	SecretKey长度错误

12.2、SDK 能力相关错误说明

code	subCode	含义
1000	1000	上下文为空
	1001	复制中
	1002	资源不存在
	1003	复制文件异常
	1004	解压文件异常
1001	2000	license 校验失败
	2001	参数异常
	2002	初始化引擎错误
1002	3000	重复加载
	3001	动作文件路径为空
	3002	读取动作数据异常
	3003	引擎加载异常
1003	4000	动作数据为加载异常
	4001	JSON 数据处理异常
	4002	drml 内容为空
	4003	drml 参数异常
	4004	解析drml 异常
1004	4000	动作数据为加载异常
	4001	JSON 数据处理异常
	4004	解析drml 异常
	4005	txt 内容为空
	4006	txt 参数异常
	4007	txt 翻译异常
	4008	txt 解密异常

云渲染直播

云渲染推流接口

🔗 接口说明

接口调用域名：<https://open.xiling.baidu.com>

使用直播 API 可以开启云端数字人渲染推流，和云渲染交互 SDK 的区别是，交互 SDK 主要用于客户端集成面向C端用户进行数字人交互的场景，而直播 API 主要用于后端对接，驱动数字人并获取数字人视频流后推送到第三方直播平台的场景。

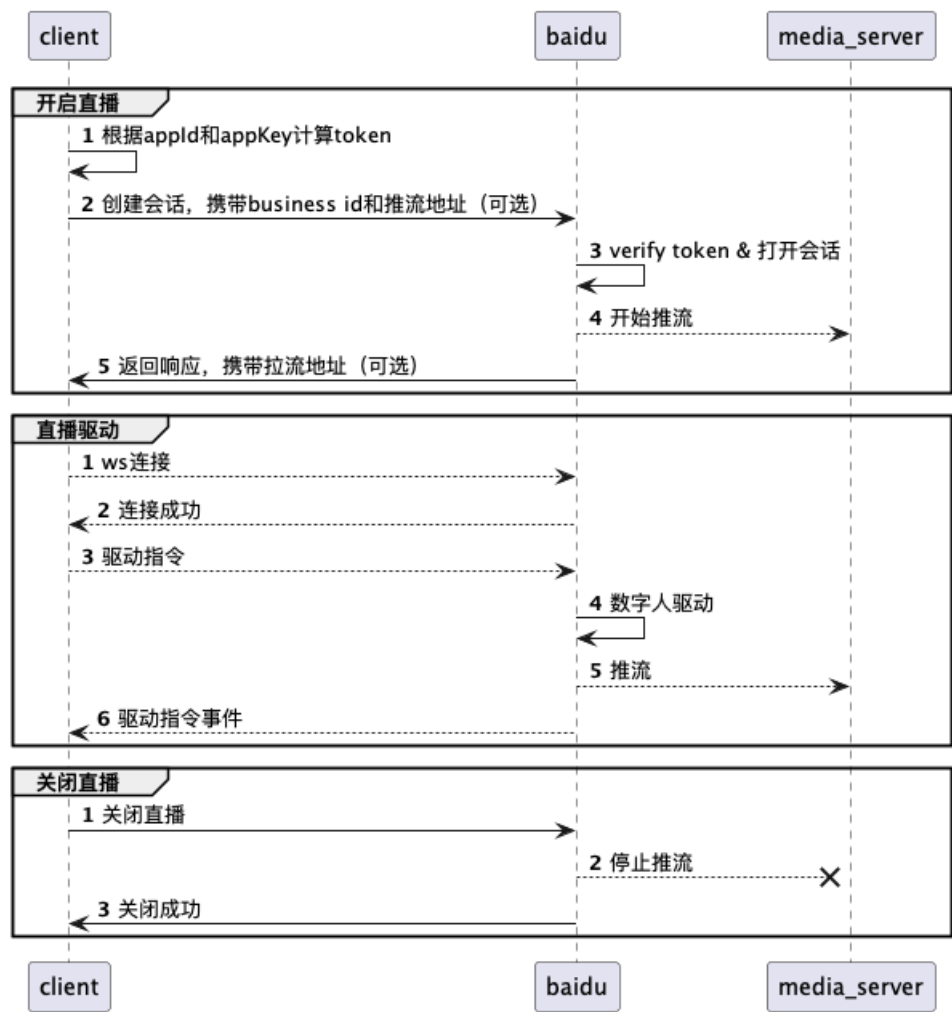
注意：这里的直播指的是数字人应用场景，我们并不提供直播分发能力，需要接入方获取数字人视频流后使用其他直播平台完成向用户的分发。

调用本接口需要在平台购买直播组件商品并绑定到应用，操作说明参考：[平台操作指引](#)

🔗 接口调用流程

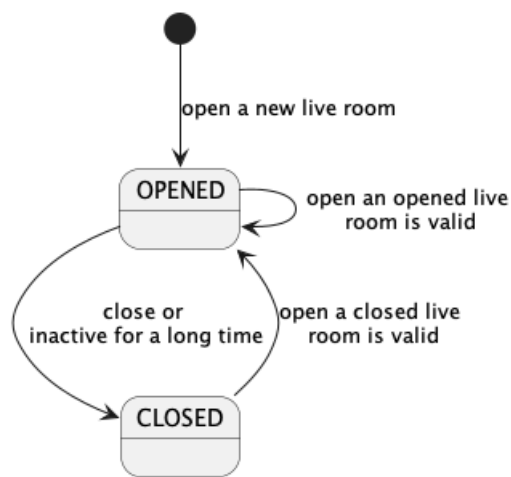
1. 通过 HTTP 请求开启、关闭直播间，查看直播间信息
2. 直播间开启后通过 Websocket 接口发送驱动数字人的指令，注意直播间开启后尽快建立 Websocket 连接，否则直播间会在

开启 1 分钟后关闭。



- client具体是客户侧的哪一个端，可以由客户选择
- 客户可自行从media_server处进行音视频流拉取

直播间状态流转



- NULL -> OPENED: 打开一个新的直播间
- OPENED -> OPENED: 打开一个正在打开的直播间是有效的，会根据新的OPEN参数重新打开直播间
- OPENED -> CLOSED: 主动关闭一个直播间或者不活跃一段时间，都会让直播间进入关闭状态
- CLOSED -> OPENED: 可以打开一个已经关闭的直播间，重新开始直播

接口列表

🔗 HTTP API

接口鉴权和通用字段说明请查阅：[接口通用说明](#)

打开直播间 request

POST /api/digitalhuman/open/v1/liveRooms **HTTP/1.1**

Content-Type: application/json

```
{
  "liveRoom": "live room", // 直播间名称，长度小于 50，同一 app 下唯一
  "resolution": { // 画面输出分辨，比例需要和人像一致，否则会出现人像比例异常
    "width": 1080,
    "height": 1920
  },
  "elements": [ // IMAGE 和 DIGITAL_HUMAN 两个 element 的顺序固定不可调整
    {
      "type": "IMAGE", // 背景图
      "image": {
        "location": { // 背景图位置和尺寸
          "top": 0, // 固定填 0
          "left": 0, // 固定填 0
          "width": 1080, // 需要和画面宽度(resolution.width)保持一致
          "height": 1920 // 需要和画面高度(resolution.height)保持一致
        },
        "url": "http://image url" // 图片地址，推荐使用 BOS（百度对象存储服务）地址，其他地址可能会因为网络原因加载图片失败导致会话打开失败
      },
    },
    {
      "type": "DIGITAL_HUMAN",
      "digitalhuman": {
        "figureId": "figure id", // 数字人形象id，只支持 2D 小样本人像，可以使用系统公共人像[2D小样本数字人]
        (AI_DH/数字人形象使用/公共形象库/2D小样本数字人.md)或定制人像[2D小样本数字人形象定制接口](AI_DH/定制服务/2D小样本数字人形象定制接口.md)
        "location": { // 数字人区域在视频里的位置，非必填参数
          "top": 0, // 数字人区域到视频顶边距离，单位：像素，默认取 0
          "left": 0, // 数字人区域到视频左边距离，单位：像素，默认取 0
          "width": 1080, // 视频中数字人区域的宽度，单位：像素
          "height": 1920 // 视频中数字人区域的高度，单位：像素
        },
        "crop": { // 裁剪参数，默认不裁剪，即取整个人像底板，非必填参数
          "top": 0, // 数字人区域到人像底板顶边距离，单位：像素，默认取 0
          "left": 0, // 数字人区域到人像底板左边距离，单位：像素，默认取 0
          "width": 1080, // 数字人区域的宽度，单位：像素
          "height": 1920 // 数字人区域的宽度，单位：像素
        },
      },
    }
  ],
  "output": {
    "mode": "BRTC", // 视频流输出模式：RTMP - 延迟较高，更适合直播场景；BRTC - 延迟较低，更适合交互场景
    "rtmp": { // mode 为 "RTMP" 时需要填写，BRTC 无需填写
      "pushUrl": "rtmp push url" // 可选，指定数字人 RTMP 推流地址，示例值:rtmp://show-pu.kaywang.cn/{app}/{steam}，如果不填写会返回平台提供的 RTMP 地址
    },
    "audioFormat": {
      "sampleRate": 16000 // 可选，会话输入音频采样率，默认16000，目前只支持16000。TEXT_REDNER或AUDIO_RENDER对应的音频采样率和该值不一致，则不会进行播报，并返回event为ERROR的消息；AUDIO_STREAM上的采样率如果和该值不一致，会进行播报，但播报声音会有问题
      "lan": "auto"
    }
  }
}
```

裁剪参数图例可以参考positionV2 参数 crop 及 location：[裁剪图例](#) response

HTTP/1.1 200 OK

```
{
  "code": 0,
  "message": "error message",
  "result": {
    "liveRoom": "live room",
    "status": "OPEN",
    "output": {
      "rtmp": {
        "pullUrl": "rtmp pull url" // output mode 为 RTMP 且请求没有提供 RTMP 推流地址(pushUrl)，系统返回内置的
        RTMP 地址供调用方进行拉流
      },
      "brtc": { // output mode 为 BRTC 时返回，参数使用参考文档直播流获取章节
        "appId": "app id", // BRTC 侧 appId，非数字人开放平台应用 ID
        "roomName": "room name", // BRTC 侧 roomName，非数字人直播间 name
        "userId": "user id", // 拉流需使用的 user id
        "userToken": "user token", // 拉流需填写的鉴权 token
        "feedId": "feed id", // 数字人推流侧的 user id
        "previewBrowseUrl": "preview url" // BRTC 模式下预览音视频 demo 网页链接
      }
    },
    "createTime": "2023-08-14 00:00:00",
    "updateTime": "2023-08-14 00:00:00"
  }
}
```

错误码 | code | 说明 | | --- | | 10006 | 参数校验不通过，根据 message 排查请求参数 | | 22001 | figure id 不存在 | | 22002 | figure id 查询失败 | | 22003 | figure id 无权限 | | 60100 | 会话开启失败，联系平台排查解决 | | 60101 | app 渲染并发路数已达上限 | | 60102 | 渲染资源不足 | | 60200 | 直播服务异常，联系平台排查解决 | | 60000 | 直播通用异常，联系平台排查解决 | | 10004 | 未识别通用异常，联系平台排查解决 |

关闭直播间 request

DELETE /api/digitalhuman/open/v1/liveRooms?liveRoom=\${liveRoom} HTTP/1.1

response

HTTP/1.1 200 OK

```
{
  "code": 0,
  "message": "error message",
  "result": {
    "liveRoom": "live room",
    "status": "CLOSE",
    "output": {
      "rtmp": {
        "pullUrl": "rtmp pull url"
      }
    },
    "createTime": "2023-08-14 00:00:00",
    "updateTime": "2023-08-14 00:00:00"
  }
}
```

查询直播间 request

```
GET /api/digitalhuman/open/v1/liveRooms?liveRoom=${liveRoom} HTTP/1.1
```

response

```
{
  "code": 0,
  "message": null,
  "requestId": "cd46a9e2-8bc5-490b-9c76-d5122ddf2b7b",
  "result": {
    "liveRoom": "test",
    "sessionId": "sl-piehqauahma2vcfzp",
    "status": "CLOSED",
    "createTime": "2023-09-04T19:43:58Z",
    "updateTime": "2023-09-05T07:00:26Z"
  }
}
```

🔗 WebSocket API

WebSocket 地址

```
wss://open.xiling.baidu.com/live/2d/ws?
appId=${app_id}&token=${app_token}&expire=${expire}&liveRoom=${live_room}
```

- app_token 和 expire 参数分别对应[接口通用说明](#)中的 Signature 和 ExpireTime。

Ping/Pong机制 百度服务端会定时向客户端发送Ping消息，客户端收到后将payload原封不动放在Pong消息里返回给百度服务端。如果服务端发现长时间没有收到Pong消息，会主动将WebSocket连接断开。

重连说明

- 在会话关闭前，ws可以重复连接，但同时只能有一条相同直播间的ws链接存在
- 重新连接ws后，会默认展示最后一次OPEN的画面
- ws断开超过一定时间（60秒），系统会自动将会话关闭，此时ws重连会失败，可通过调用打开接口重新打开后进行重连
- 在ws断开过程中产生的事件会丢失，无法发送到客户端，但调用方可通过Ready消息内容来做一些业务判断

通用业务消息格式

- 上行代表调用方发送给百度服务器，下行代表百度服务器发送给调用方
- 下行的消息的id和上行的id进行对应，方便客户端进行判断处理，**注意：上行消息 id 需为大于 0 的整数，递增且全局不重复。**
- 上下行消息的body为json string格式，后续各类型的消息结构就是body的json格式

up

```
{
  "id": 1, // 大于 0 的整数，递增不能重复
  "type": "msg type",
  "body": "msg body in json"
}
```

down

```
{
  "id": 1, // 大于0的整数
  "type": "msg type",
  "code": 0, // 0 - 成功；非0 - 失败
  "message": "error msg" // 异常信息
  "body": "msg body in json"
}
```

服务端主动断开异常码 | code | 说明 | | --- | --- | | 4000 | 直播房间不存在 | | 4001 | 直播房间未打开 | | 4100 | 渲染流打开超时 | | 4101 | 渲染流打开被打断 | | 4102 | 渲染流打开异常 | | 4103 | 渲染流异常 | | 4104 | 渲染流关闭 | | 4105 | 渲染流重复 | | 4200 | ping/pong超时 | | 4300 | 用户输入异常 |

Ready type = READY

ws连接成功后，由后端返回的状态信息，收到该消息表明整个链路完全建立，可以开始发送驱动指令

down

```
{
  "status": "IDLE", // 后端状态：IDLE - 空闲；SPEAKING - 播报中
  "id": 1 // 最近一个正在播报（SPEAKING）或已经播报完（IDLE）的RENDER指令的 id
}
```

- Ready消息外层的 id 为 0
- 如果status为SPEAKING，但是在Ready之前有event为COMPLETE的消息返回，那么SPEAKING消息无意义，此时调用方可根据COMPLETE消息进行业务判断

TextRender type = TEXT_RENDER

文本驱动

up

```
{
  "textId": "text id", // 自定义id，便于问题排查
  "text": "text", // 播报文本，长度不能超过 1000;当 lan 参数非'auto'时，请自行确认文本与所选语言是否一致，如不一致合成成功后会导致发音异常。
  "tts": {
    "per": "5116", // 音色id，公共音色列表参考：[公共音色库](AI_DH/数字人形象使用/公共音色库.md)，定制音色参考：[音色克隆和查询接口](AI_DH/定制服务/音色克隆和查询接口.md)
    "lan": "Chinese", // 仅有LITE版克隆音色及公共音色库中的支持多语言的音色支持该参数
    "spd": "5", // 语速：5是正常值，0-15的取值范围。越大语速越快，默认值5（字面量需要是整数）
    "vol": "5", // 音量：5是正常值，0-15的取值范围，越大音量越大，默认值5（字面量需要是整数）
    "pit": "5" // 语调：5是正常值，0-15的取值范围，越大声音越尖，默认值5（字面量需要是整数）
  },
  "insert": false, // 为 true 则插入到当前播报中 TEXT_RENDER 的未播报文本最近的一个标点符号后播报或当前播报中 AUDIO_RENDER 的音频后播报，默认 false，多个 TEXT_RENDER 消息排队播报
  "delayed": true // insert 为 false 时生效，为 true 则会在播报前延迟约 1 秒再播报，模拟真人播报的停顿
}
```

down

```
{
  "event": "START" // START, COMPLETE, ERROR, INTERRUPTED
}
```

AudioRender type = AUDIO_RENDER

单个音频文件驱动

up

```
{
  "audioid": "audio id", // 自定义id, 便于问题排查
  "url": "audio url", // 音频需要满足: WAV格式, 16k 采样率, 通道数 1, 位深 16 bit, WAV header 44 字节, 推荐使用
  BOS (百度对象存储服务) 地址, 其他地址可能会因为网络原因加载失败导致驱动失败
  "insert": false, // 为 true 则插入到当前播报中 TEXT_RENDER 的未播报文本最近的一个标点符号后播报或当前播报中
  AUDIO_RENDER 的音频后播报, 默认 false, 多个 AUDIO_RENDER 消息排队播报
  "delayed": true // insert 为 false 时生效, 为 true 则会在播报前延迟约 1 秒再播报, 模拟真人播报的停顿
}
```

可以使用我们提供的素材库接口上传音频, 进行转码, 获取URL (即将上线)

down

```
{
  "event": "START" // START, COMPLETE, ERROR, INTERRUPTED
}
```

AudioStream type = AUDIO_STREAM

音频流驱动, 音频需要保证PCM格式, 通道数1, 位深16bit。默认采样率16000, 如果是其他采样率, 需要在上行START消息中指定。

客户端先发上行START消息, 待收到服务端下发的event为READY的消息后代表可以开始发送音频数据。

之后客户端可以通过websocket上行的二进制数据通道发送音频数据, 当音频发送结束时, 客户端发送上行COMPLETE事件。

在一对START和COMPLETE之间, 除了INTERRUPT之外的其他文本类型的ws消息都会被认为是异常请求。

up

```
{
  "streamId": "stream id", // 自定义id, 便于问题排查
  "event": "START", // START, COMPLETE
  "sampleRate": 16000 // 默认 16000
}
```

- 如果想要确保插入在当前播报的一句话之后, 请在这句话播报结束前3秒之前发送插入指令

down

```
{
  "streamId": "stream id", // stream id
  "event": "START" // READY, START, COMPLETE, ERROR, INTERRUPTED
}
```

up binary 调用方在收到下行成功的START后, 可以通过ws的binary数据通道发送流式音频, 该音频会用于实时驱动数字人:

- 调用方需要确保发送速率, 避免卡顿
- 单次上行binary数据最大不超过256KB

Interrupt type = INTERRUPT

打断播报, 包括清空buffer中还没有播报的内容:

- INTERRUPT会有一个对应id和类型的下行消息，代表INTERRUPT被处理了
- 被打断的播报内容会触发event为INTERRUPTED的下行事件消息，对应的外层id为触发播报内容的驱动指令外层id

直播流获取

RTMP 打开房间请求体中 `output.mode == RTMP`

可直接通过响应中的RTMP链接获取直播流。RTMP 流存在秒级的延迟，如果业务上需要准确对齐视频内容建议使用 BRTC 流。

BRTC 打开房间请求体中 `output.mode == BRTC`

输出采用BRTC的方式，时延较低，在这种模式下，获取输出流需要接入百度RTC的SDK，请参考 [实时音视频介绍](#)

联系我们

1 业务咨询

- 客服电话：400-920-8999
- 提交工单：<https://console.bce.baidu.com/support/#/ticket/#/ticket/create>
产品选择“曦灵数字人开放平台”，描述具体问题
- 合作咨询：<https://cloud.baidu.com/survey/assembly.html>

2 帮助文档

帮助文档：[产品概述](#)

3 官方助手



曦灵数字人定制组件克隆协议

1、 您使用平台声音克隆、照片数字人克隆、2D小样本数字人克隆功能时，承诺您所提供/上传/输入/发布的知识文档、数据、文本、文件、资料、链接等全部信息(包括但不限于上传的任何声音、图片、视频)内容来源合法合规，不侵犯任何第三方的合法权益(包括但不限于包括声音权、肖像权在内的人格权、著作权、专利权、商标权等)；您保证对其上传的所有内容拥有完整且合法的所有权、使用权、知识产权，对于需要获得第三方明示同意或授权的已经取得相应证明，对于您使用我们提供的克隆功能，所编辑的非您本人的人脸、人声等生物识别信息，您应依法告知被编辑的个人，并取得其单独同意，否则您承担所有不利后果。若前述信息内容中包含终端用户或任何第三方的相关个人信息，您保证您所收集和提供上述个人信息的行为已遵循合法、正当、必要的原则，并承诺，我们基于您的指令所分析、处理前述信息内容以及生成新的信息内容不会侵犯任何第三方的合法权益。在法律允许的范围内，我们可对前述信息内容进行显示或使用，如上述内容违反了任何法律或平台协议或政

策，我们有权进行删除。

2、您应自行对其提供的声音、图片、视频等信息资料、利用我们提供的服务所从事的视频制作、直播、对话服务等活动、通过我们提供的产品进行的活动所引起的一切后果承担全部责任，我们对此不承担任何责任。

3、不得利用本公司提供的资源和服务进行或允许第三方进行如下活动：

- 1) 利用本公司提供的资源和服务，从事窃取、泄露、转让、滥用、篡改或破坏包括本公司在内的数据、商业秘密和/或信息，包括但不限于 TTS 数据、个人信息、生物识别数据等；
- 2) 使用上述合法或非法获取的数据、商业秘密和/或信息，以直接或间接的方式创建、训练或改进相似或存在竞争的产品和服务；
- 3) 对本服务进行逆向工程、拆卸、反向汇编、反编译、编译，或以其他方式试图发现或取得本服务或第三方服务的源代码、系统的源代码或底层组件、底层技术、算法、模型；
- 4) 将有关干扰、破坏或限制任何计算机软件、硬件或通讯设备功能的软件病毒、木马、蠕虫、逻辑炸弹或其他恶意或有害的计算机代码、档案和程序之资料，加以上载、发布、发送电子邮件或以其他方式传送；
- 5) 非法入侵本服务服务器、干扰网络正常功能、窃取网络数据、篡改本服务之相关代码或其他损害本服务的行为；或提供专门用于从事侵入网络、干扰网络正常功能及防护措施、窃取网络数据等危害网络安全活动的程序、工具；
- 6) 企图规避平台的安全设置或网络系统，包括获取使用者不应获取的数据，登录未经明确授权的服务器或账号，或采用运行端口扫描等方式窥探其他网络的安全措施。
- 7) 未经我们授权，非法转售或对外提供本服务，或以商业化目的使用本服务。
- 8) 破坏、绕开账号校验机制。
- 9) 其他非法克隆TTS 数据、不正当竞争、知识产权侵权等行为。

4、您充分了解并同意，本产品仅为用户提供本协议规定的服务内容，无法代替用户进行信息判断和行为决策。相应地，您应了解，使用本产品服务可能发生来自他人非法或不当行为(或信息)的风险，您应自行判断及行动，并自行承担相应的风险。

5、您使用声音克隆、照片数字人克隆、2D 数字人克隆产品及相关服务，以及输入、输出/生成、制作、上传、发布、传播的信息应自觉遵守适用法律、行政法规以及规范性文件的要求，遵守公共秩序，尊重社会公德和伦理道德、坚持社会主义核心价值观、维护国家利益、尊重他人合法权益、道德风尚和信息真实性等要求，否则我们有权立即采取相应处理措施(包括但不限于:注销账号 等)。您承诺在使用克隆产品及相关服务时，不得也不允许或协助任何人利用本服务采取以下行为：

- (1)反对宪法所确定的基本原则。
 - (2)危害国家安全，泄露国家秘密，颠覆国家政权，推翻社会主义制度、损害国家形象、煽动分裂国家、破坏国家统一的。
 - (3)损害国家荣誉和利益。
 - (4)侵害社会公共利益、扰乱经济和社会秩序。
 - (5)煽动、宣扬民族仇恨、民族歧视，破坏民族团结。
 - (6)侮辱、滥用英烈形象，否定英烈事迹，美化粉饰侵略战争行为的。
 - (7)破坏国家宗教政策，宣扬邪教和封建迷信。
 - (8)散布谣言、污言秽语，扰乱社会秩序或公序良俗，破坏社会稳定。
 - (9)宣扬淫秽、色情、赌博、暴力、凶杀、恐怖主义、极端主义或者教唆犯罪。
 - (10)煽动非法集会、结社、游行、示威、聚众扰乱社会秩序。
 - (11)侮辱、诋毁、诽谤或以其他方式伤害他人，泄露他人隐私、个人信息或资料，侵害他人名誉权、肖像权、声音权、隐私权、知识产权和其他合法权益。
 - (12)侵害未成年人合法权益或者损害未成年人身心健康的。
 - (13)破坏国际关系以及国际和平稳定。
 - (14)人身攻击及辱骂他人。
 - (15)针对以下群体发表诅咒、歧视、仇恨、漠视生命尊严、偏见等性质的言论，造成或可能造成该人或另一人的身体或心理伤害，包括但不限于民事歧视、宗教歧视、国籍歧视、性别歧视、性取向歧视、年龄歧视、职业歧视、健康歧视、地域歧视。该等群体包括:不同特定国籍、地域、性别、性别认同、性倾向、种族、民族、健康状况、职业、年龄、信仰、残障群体等。
 - (16)包含恐怖、暴力血腥、高危险性、危害表演者自身或他人身心健康内容的。
 - (17)对他人进行诅咒、恐吓或威胁。
 - (18)制作、复制、发布、传播虚假信息。
 - (19)制作、复制、发布、传播电信网络诈骗信息、开展涉诈活动。
 - (20)含有法律、行政法规、政策、公序良俗禁止的，干扰本平台运营或侵犯其他用户或第三方合法权益内容其他内容的信息。
- 您可向我们投诉、举报各类违法违规行为、违法传播活动、违法有害信息等内容，我们将及时受理和处理您的投诉举报，以共同营造风清气正 的网络空间。

4、您还应遵守本软件/平台的《服务协议》、《隐私政策》(包括其不时地更新版本)的约定。

5、如您违反前述规定，我们有权随时限制或拒绝您使用账号，暂停或关闭服务，甚至注销您的账号。

6、因您违反法律法规、《服务协议》或本使用规范造成处罚、违约或侵犯第三方合法权益等的，由您自行解决。如果给我们造成损失，我们有权向您追偿其因此所遭受的全部损失(包括其直接损失、间接损失、预期利益损失等)，我们因维权所支出的合理费用(包括但不限于律师费、差旅费、公证费、诉讼费等)均由您承担。

MCP服务

数字人MCP服务使用说明

简介

百度智能云曦灵数字人开放平台，现在已经全面适配MCP协议。欢迎各位创作者接入体验。

数字人开放平台面向企业及个人开发者，提供高性能、易集成和多样化的数字人组件能力，满足多类型数字人形象及声音定制、视频合成、交互对话、直播等应用需要，支持公有云调用与私有化部署。帮助开发者快速集成数字人能力，助力企业打造全场景应用。智能集成，轻松搞定。

更多更新数字人相关功能，欢迎访问 [百度智能云曦灵数字人开放平台](#)。

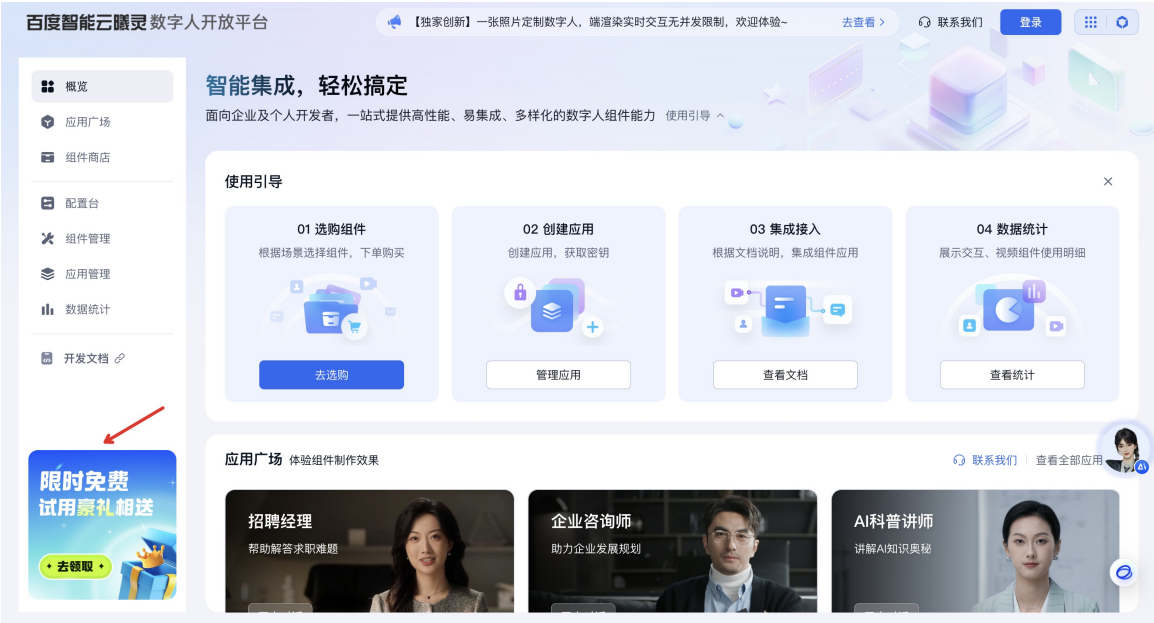
版本	更新说明	更新时间
0.1.0	现已全面支持 SSE及STDIO模式(python)，提供包含数字人形象定制，声音定制，视频合成等多个工具。	2025-05-20

使用说明

1. 接入前准备

1.1 领取试用额度

- 登录 [百度智能云曦灵数字人开放平台](#) 点击左下角



- 进入 [组件管理](#) 查看获取的组件额度



1.2 获取API Key和Secret Key

- 进入 [应用管理](#) 配置需要使用的组件



- 创建完成后即可获取 AppID 和 AppKey



1.3 MCP配置

- STUDIO模式配置参数如下：

支持Python调用。使用STUDIO模式时，在终端配置时需要传入 AppID 和 AppKey。

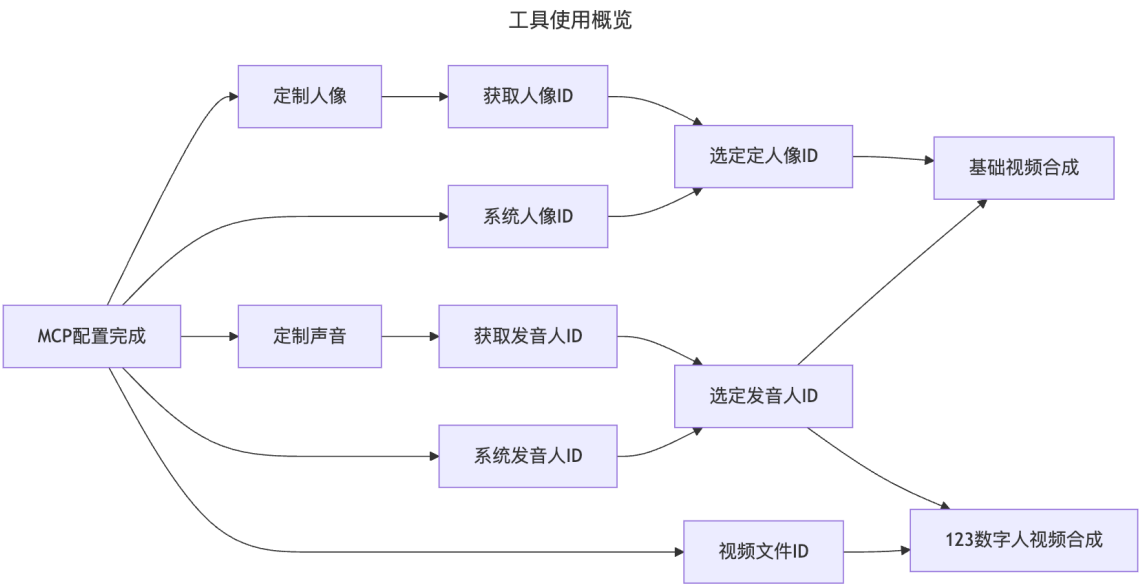
```
{
  "mcpServers": {
    "DH_MCP": {
      "command": "uvx",
      "args": ["mcp-server-baidu-digitalhuman"],
      "env": {
        "DH_API_AK": "<YOUR AK>",
        "DH_API_SK": "<YOUR SK>"
      }
    }
  }
}
```

- SSE模式配置参数如下：

使用SSE模式时，需要使用 AppID 和 AppKey 生成鉴权参数，可以使用 [数字人鉴权工具](#) 生成。也可以[参考文档](#) 生成。

```
{
  "mcpServers": {
    "DH_MCP": {
      "url": "https://open.xiling.baidu.com/mcp/sse?token=xxx"
    }
  }
}
```

2. 接入方式



2.1 功能总览

● 功能总览

我们提供了多种工具，满足不同场景下的需求。让您在大模型中快速集成数字人服务，轻松打造数字人应用。如您对数字人有更多期望，希望更加深入集成数字人服务，欢迎访问 [百度智能云曦灵数字人开放平台](#) 联系我们。更多功能也将在MCP中逐步开放，敬请关注。

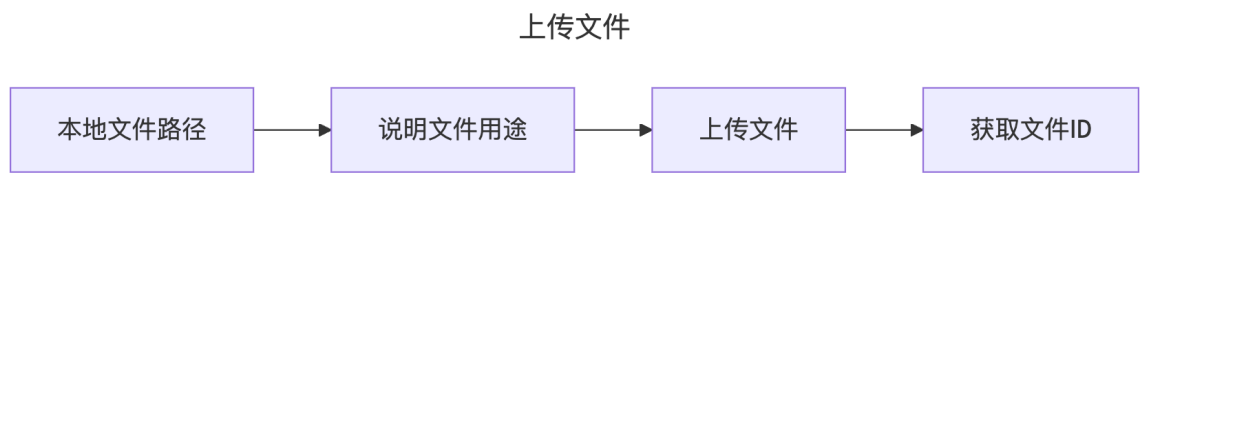
功能	功能说明	包含工具	支持模式
文件上传	根据业务类型上传所需要的文件。	uploadFiles	STDIO
音色查询	查询可用的系统发音人ID。	getVoices	SSE & STDIO
人像查询	查询可用的系统发音人ID。	getFigures	SSE & STDIO
声音克隆	根据上传音频生成音色，可用于语音合成及视频制作。	- generateVoiceClone - getVoiceCloneStatus	SSE & STDIO
2D小样本数字人	根据上传真人录制的视频生成数字人像，仅可用于基础视频制作，数字人使用通用口型驱动。	- generateLite2dGeneralVideo - getLite2dGeneralStatus	SSE & STDIO
数字人视频合成	根据所选数字人像及音色，生成数字人视频	- generateDhVideo - getDhVideoStatus	SSE & STDIO
123数字人视频	根据真人录制的视频及选定音色，无需人像生成，直接生产对应的数字人视频	- generateDh123Video - getDh123VideoStatus	SSE & STDIO
语音合成	根据提供的文本内容及选定音色，无需生成视频，生产对应的音频	- generateText2Audiooo - getText2AudioStatus	SSE & STDIO

2.2 文件上传

- 功能说明：平台支持上传音频、视频文件，用于后续的声音克隆，数字人制作，123数字人视频制作等。

该功能暂只支持STDIO模式。SSE模式中，如需获取文件ID，可先参考[文件上传接口](#) 使用接口实现。我们将在后续的更新中对该工具进行升级。

- 使用流程：



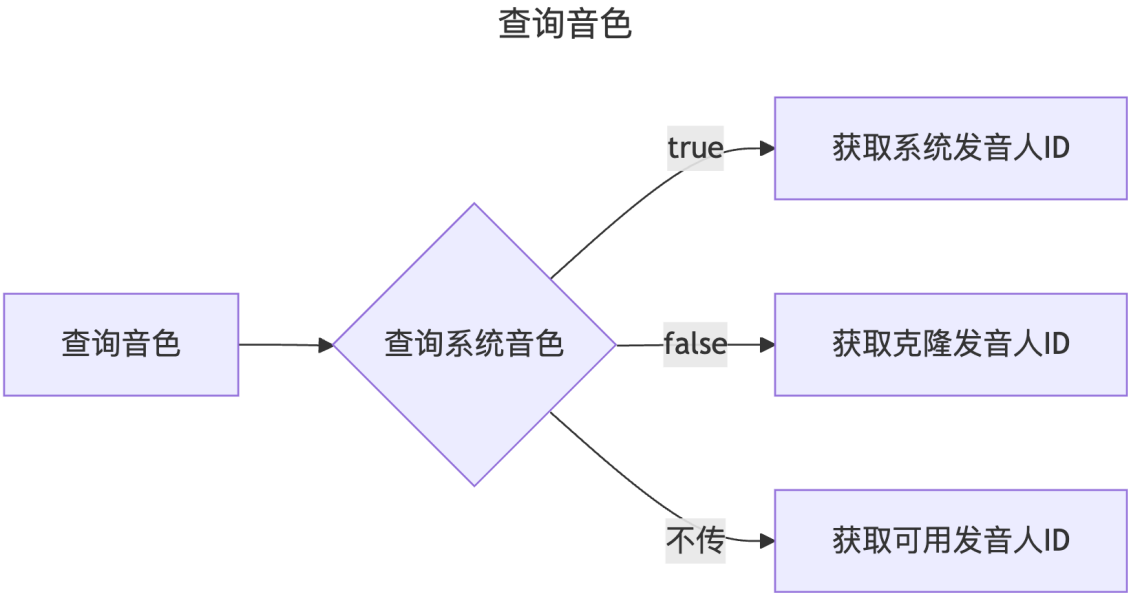
- 示例提示词： 上传test.mp3这个文件用于声音克隆，文件在C：/Users/username/Desktop/test.mp3。

- 工具详情：

工具名称	工具描述	输入参数	输出内容
uploadFiles	根据业务类型上传所需要的文件。	- file：需要上传的文件 - providerType：使用这个文件的业务类型，目前仅限于“2D小样本数字人制作”，“声音克隆”，“123数字人视频制作”三种业务类型。 - sourceFileName：上传的文件名，必须填写正确的文件名称及后缀，比如：test.mp3。	- fileId：文件ID - fileName：上传的文件名

2.3 音色查询

- 功能说明：查询可用的系统发音人ID。
- 使用流程：

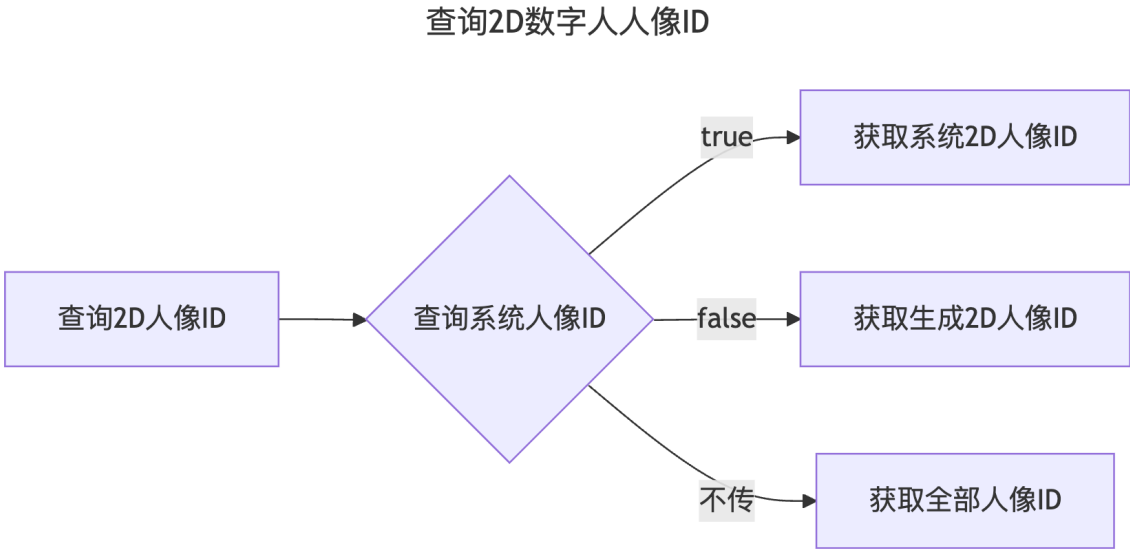


- 示例提示词： 我之前克隆过哪些声音？ 我想用一个二十岁左右温柔小姐姐的声音。
- 工具详情：

工具名称	工具描述	输入参数	输出内容
getVoices	查询可用的发音人ID。	• isSystem：“true”查询系统发音人ID，“false”查询克隆发音人ID，不传任何值则为查询可用发音人ID	• perId:发音人ID • name：发音人名称 • describe：音色特点的描述 • gender：性别 • systemProvided：是否是系统音色

2.4 人像查询

- 功能说明：查询可用的2D数字人人像ID。
- 使用流程：

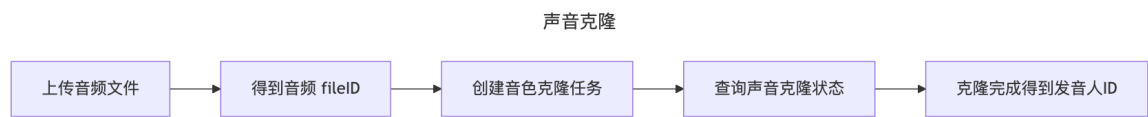


- 示例提示词： 我之前生成过哪些人像？ 有哪些可用的人像？
- 工具详情：

工具名称	工具描述	输入参数	输出内容
getFigures	查询可用人像ID。	• isSystem：“true”查询系统人像ID，“false”查询生成人像ID，不传任何值则为查询可用人像ID	• figureId:2D人像ID • name：2D人像名称 • gender：性别 • systemProvided：是否是系统音色

2.5 声音克隆

- 功能说明：根据上传音频生成音色，可用于语音合成及视频制作。
- 使用流程：



- 示例提示词： 用文件id为xxx的音频文件克隆声音。命名为“zhangsan”，是一个三十岁左右中年男性的音色，用“这个是我克隆的声音”这段文本试听一下。 查一下id为xxx的声音克隆好了没有。
- 工具详情：

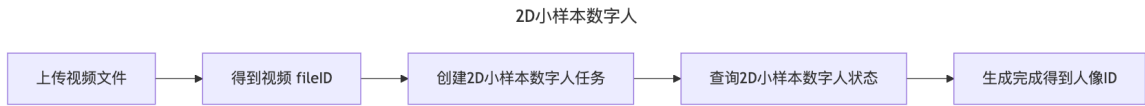
工具名称	工具描述	输入参数	输出内容
generateVoiceClone	根据上传音频生成音色，可用于语音合成及视频制作。	• name：对克隆音色的命名，长度不超过 50 • gender：发音人的性别 • describe：对克隆音色的描述，长度不超过 100 • uploadAudioId：用于克隆音色的音频对应的文件ID • example：用于试听的文本，长度不超过 100	• perId；被克隆音色的发音人ID
getVoiceCloneStatus	根据声音克隆任务的发音人ID，查询该任务目前的状态。	• isSuccess：是否只查询克隆成功的任务 (true: 只查询成功的任务， false: 查询全部克隆任务) • perId；：查询指定发音人ID的任务	• perId；被克隆音色的发音人ID • name：发音人的名称 • describe：对克隆音色的描述 • exampleText：用于试听的文本 • exampleAudioUrl：使用试听的文本合成的音频文件的链接 • status：当前任务的状态，PREPARING(准备中), CLONING(克隆中), SUCCESS(克隆成功), FAIL(克隆失败) • reason：如果克隆失败，则此处会描述失败原因 • gender：被克隆音色的发音人的性别

2.6 2D小样本数字人

- 功能说明：根据上传真人录制的视频生成数字人像，仅可用于基础视频制作，数字人使用通用口型驱动。

暂时只支持使用透明背景的webm视频来生成数字人像。

- 使用流程：



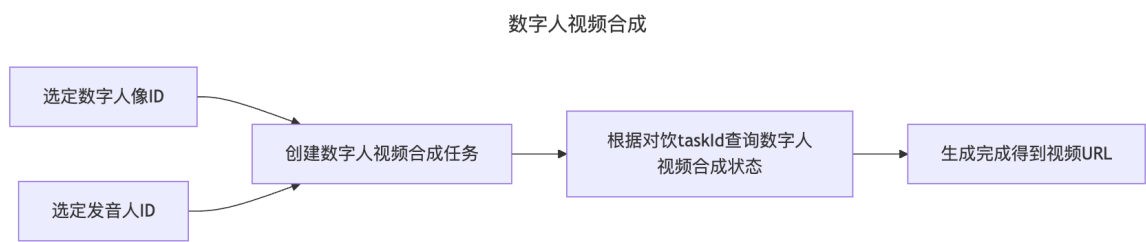
- 示例提示词： 用fileid为xxx的视频文件，生成数字人，命名为“zhangsan”，是个男生的形象。 查一下id为xxx的数字人好了没有。 我可以用哪些人像。

- 工具详情：

工具名称	工具描述	输入参数	输出内容
generateLite2dGeneralVideo	根据上传真人录制的视频生成数字人像，仅可用于基础视频制作，数字人使用通用口型驱动。	- name：对生成数字人像的命名，长度不超过50 - gender：数字人的性别 - keepBackground：是否保留视频背景，true为保留，false为去除，默认为false - templateVideoId：用于生成数字人像的视频对应的文件ID	figureId：根据上传真人录制的视频生成的数字人像ID
getLite2dGeneralStatus	- 查询数字人像的生成的进度 - 也可以用于查询有哪些可用的系统2D人像。	- figureId：指定人像ID查询，为空则查询该账号下所有人像 - systemFigure：查询平台的公共人像，为空：查询全部，true：返回平台公共人像，false：返回定制人像 - trainSuccess：是否查询训练完成，状态可用的人像（为空：不进行过滤，true：只返回可用人像（平台公共人像和训练成功状态的定制人像），false：只返回排队中、训练中或训练失败的定制人像） - pageNo：页码，默认为1 - pageSize：每页容量，默认10	- figureId：根据上传真人录制的视频生成的数字人像ID - name：对生成数字人像的命名，长度不超过50 - gender：数字人的性别 - keepBackground：是否保留视频背景，true为保留，false为去除，默认为false - status：状态（LINE_UP（排队中），GENERATING（训练中），SUCCESS（训练成功），FAILED（训练失败）） - failedCode：失败错误码 - failedMessage：制作失败原因

2.7 数字人视频合成

- 功能说明：根据所选数字人像及音色，生成数字人视频
- 使用流程：

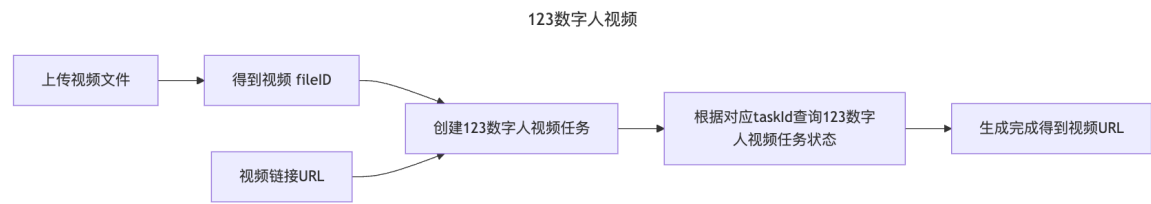


- 示例提示词： 用数字人像ID为xxx，发音人ID为yyy的音色，视频的内容是“大家好，我是数字人播报的内容”，使用横屏全身的机位，视频背景用“<https://digital-human-material.bj.bcebos.com/-%5B%3B%4046f6cc1e.png>”，开启自动添加动作，开启字幕，生成一个1080P的数字人视频。 查一下taskid为xxx的数字人视频好了没有。
- 工具详情：

工具名称	工具描述	输入参数	输出内容
generateDhVideo	根据所选数字人像及音色，生成数字人视频。	- figureId：数字人像ID - driveType：驱动数字人的数据类型，支持文本驱动或者音频驱动 - text：驱动类型为文本驱动时，必填的视频内容，长度不超过20000 - person：驱动类型为文本驱动时，必填的发音人ID - inputAudioUrl：驱动类型为音频驱动时，必填的音频链接URL - width：输出视频分辨率的宽 - hight：输出视频分辨率的高 - cameraId：系统人像的机位设置，0: 横屏半身, 1:竖屏半身, 2: 横屏全身, 3: 竖屏全身 - enabled：是否开启字幕，true开启字幕，默认false不开启。 - backgroundImageUrl：背景图片URL - autoAnimoji：系统人像自动添加动作，true为自动添加，默认为false不添加	taskId：当前视频合成的任务ID
getDhVideoStatus	查询数字人视频合成进度。	taskId：当前视频合成的任务ID	- taskId：当前视频合成的任务ID - status：状态：SUBMIT（已提交待合成）,GENERATING（合成中）,SUCCESS（合成成功）,FAILED（合成失败） - failedCode：失败错误码 - failedMessage：制作失败原因 - videoUrl：任务ID对应的成功合成的视频文件地址，文件会保存 7 天

2.8 123数字人视频合成

- 功能说明：根据真人录制的视频及选定音色，无需人像生成，直接生产对应的数字人视频
- 使用流程：

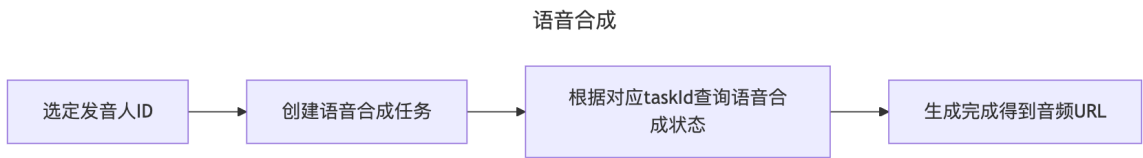


- 示例提示词：用fileid为xxx的视频文件，发音人ID为yyy的音色，视频的内容是“大家好，我是数字人播报的内容”，生成一个数字人视频。 视频的地址是https://open-api-test.bj.bcebos.com/ae870923-2a3b-4d5e-b6a2-e44b4025647220250417_163529_trim.mp4，用发音人ID为yyy的音色，视频的内容是“大家好，我是数字人播报的内容”，生成一个数字人视频。 查一下taskId为xxx的123数字人视频好了没有。
- 工具详情：

工具名称	工具描述	输入参数	输出内容
generateDh123Video	根据真人录制的视频及选定音色，无需人像生成，直接成一个数字人视频。	• templateVideoId：用于生成数字人视频的视频对应的文件ID或URL • driveType：驱动数字人的数据类型，支持文本驱动或者音频驱动 • text：驱动类型为文本驱动时，必填的视频内容，长度不超过20000 • person：驱动类型为文本驱动时，必填的发音人ID • inputAudioUrl：驱动类型为音频驱动时，必填的音频链接URL	• taskId：当前视频合成的任务ID
getDh123VideoStatus	查询123数字人视频合成进度。	• taskId：当前视频合成的任务ID	• taskId：当前视频合成的任务ID • status：状态：SUBMIT（已提交待合成）,GENERATING（合成中）,SUCCESS（合成成功）,FAILED（合成失败） • failedCode：失败错误码 • failedMessage：制作失败原因 • videoUrl：任务ID对应的成功合成的视频文件地址，文件会保存 7 天

2.9 语音合成

- 功能说明：根据提供的文本内容及选定音色，无需生成视频，生产对应的音频。
- 使用流程：



- 示例提示词：用发音人ID为xxx的音色，内容是“大家好，我是数字人播报的内容”，生成音频。· 查一下taskId为xxx的语音合成好了没有。·

工具名称	工具描述	输入参数	输出内容
generate Text2Audio	根根据提供的文本内容及选定音色，无需生成视频，生产对应的音频。	• text：必填的文本内容，长度不超过2000 • person：必填的发音人ID	• taskId：当前音频合成的任务ID
getText2AudioStatus	查询音频合成进度。	• taskId：当前视频合成的任务ID	• status：状态：SUBMIT（已提交待合成）,GENERATING（合成中）,SUCCESS（合成成功）,FAILED（合成失败） • failedCode：失败错误码 • failedMessage：制作失败原因 • audioUrl：任务ID对应的成功合成的音频文件地址，文件会保存 7 天

🔗 3. 使用声明

当您使用以上工具前，请先阅读 [曦灵数字人定制组件克隆协议](#)。当您使用以上工具时，即表示您同意该协议。

🔗 联系我们

如果您有任何问题或建议，请随时联系我们。您可以通过以下方式联系我们：

- 客服电话：400-920-8999
- 合作咨询：[百度智能云曦灵数字人开放平台咨询](#)
- 问题工单：[创建工单](#)
- 官方助手：

