

Tencent 腾讯 | CSIG | 腾讯云智能

3D数字人AI驱动技术及产品介绍

数字人技术产品经理：陈伟杰



目录

3D 数字人AI驱动脸部动画

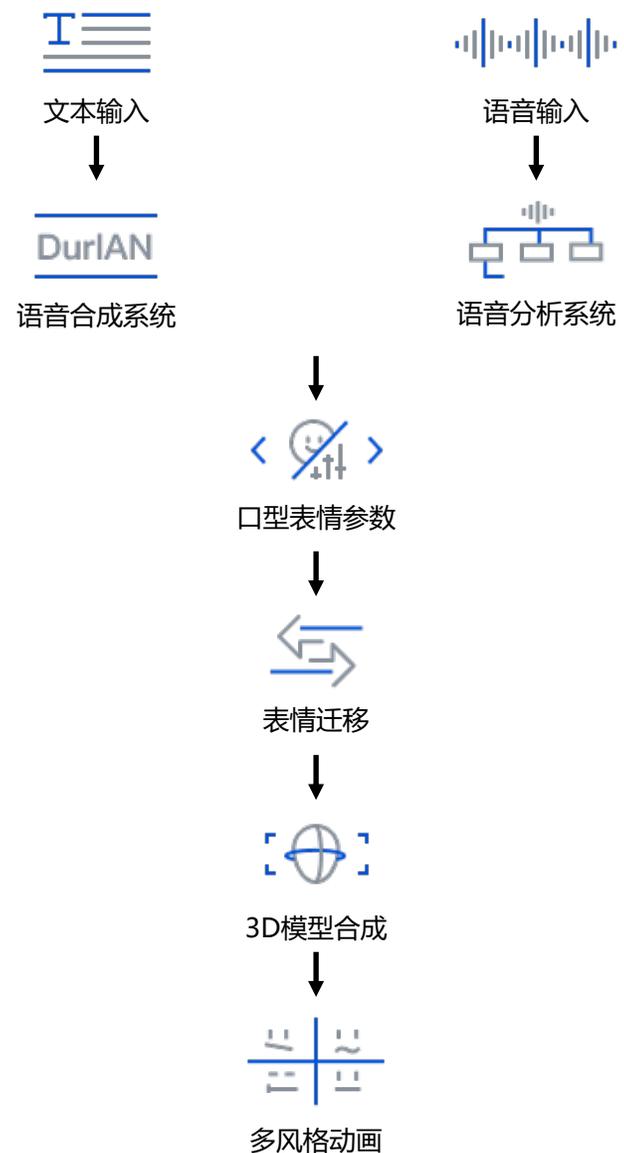
3D 数字人AI驱动身体动画

应用案例

平台及管线价值

3D数字人AI驱动脸部动画

3D数字人口型及表情AI驱动技术可以通过**语音或文本**输入合成**口型准确、表情生动**的3D角色模型脸部动画，能有效提升3D数字人在游戏、动画等领域的角色面部动画制作效率及品质。



3D数字人AI驱动脸部动画 - 平台化管线化解决方案

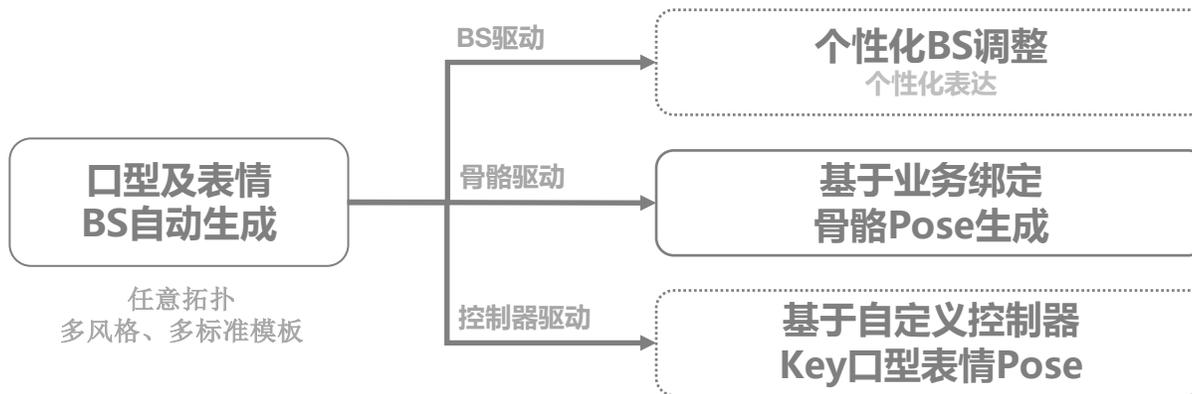
业务：资产准备

平台：资产制作

DCC工具：动画生成

脸部资产

数字人中立脸

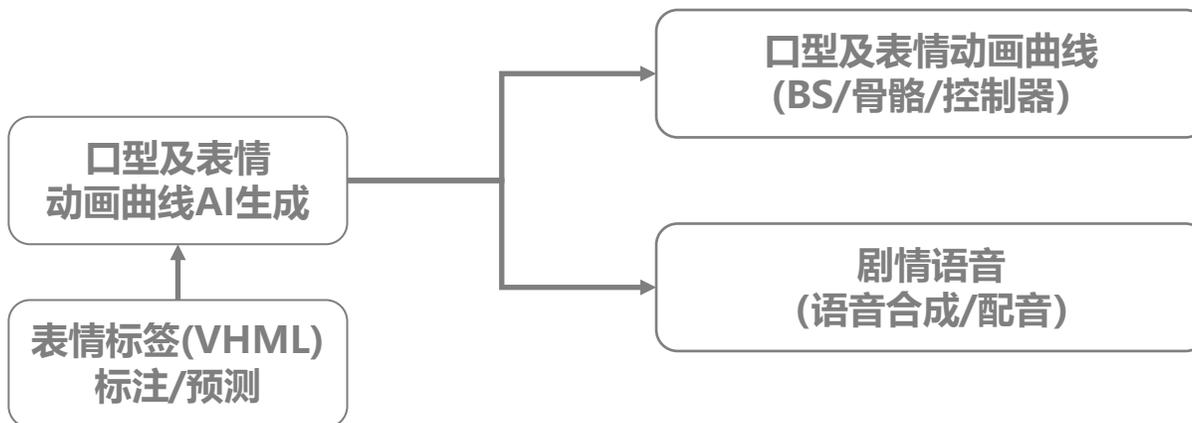


动画曲线

剧情台本



文本/语音



3D数字人AI驱动脸部动画 - 平台化管线化解决方案

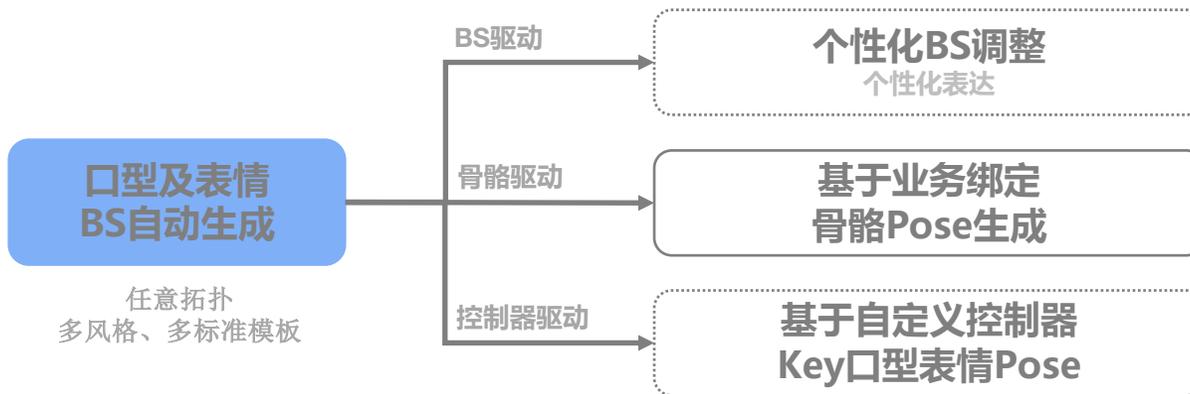
业务：资产准备

平台：资产制作

DCC工具：动画生成

脸部资产

数字人中立脸



Maya工具

AutoRig 工具
自定义表情设置
细粒度动画调整

BS动画压缩
(optional)

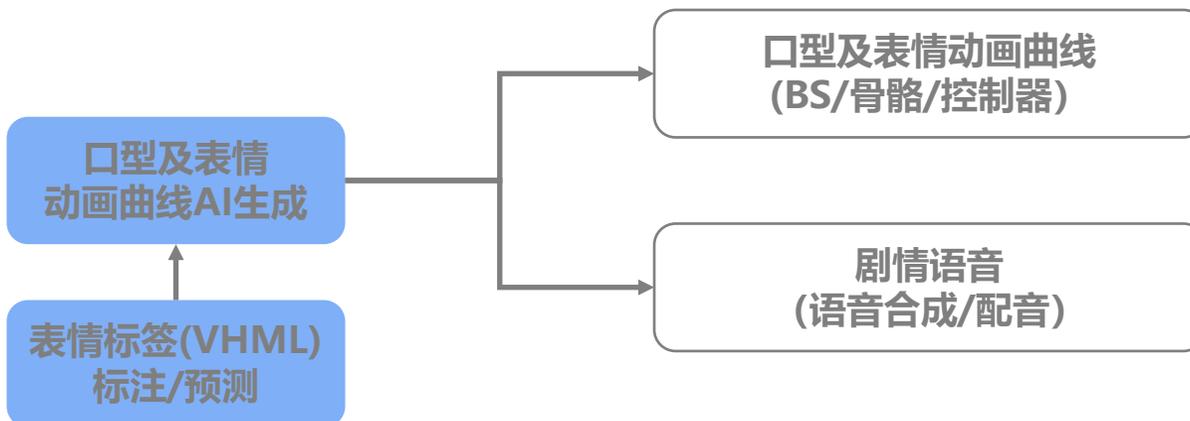
降低资源
统一制作管线

动画曲线

剧情台本



文本/语音



引擎

In-game/离线
动画生成

3D数字人AI驱动脸部动画 – 口型及表情BS/骨骼Pose生成

口型及表情AI驱动服务 > 口型及表情参考模板生成

BS自动生成 FBX骨骼pose文件自动生成

* 模型标准:

VISM16  VISM16+AU20  VISM16+ARKit  ARKit 

* 选择模板:



写实男性角色



写实女性角色



美型男性角色



美型女性角色

启用BS模型增强:

关

* 上传文件:

📁 业务中立脸模型 *

obj格式 10M 以内

📁 业务张嘴模型 *

obj格式 10M 以内

📁 Wrap 后的张嘴模型 *

obj格式 10M 以内

 男性标准脸

 女性标准脸

确认提交

- 基于模板BS自动生成
- 基于自定义骨骼的pose拟合
- 兼容自研及业界主流绑定标准BS模板
 - 自研BS标准 VISM16+AU20
 - 业界常见BS标准: ARKit BS
 - 混合方案: VISM16+ARKit
- 多风格模板: 写实 美型/卡通 男女
- 支持任意拓扑、自定义骨骼

3D数字人AI驱动脸部动画 – 脸部动画曲线AI生成

Speech2Face

驱动配置 ⓘ

* 帧率:

30fps 60fps

* 输出参数类型:

VISM16 ⓘ VISM16+AU20 ⓘ VISM16+ARKit ⓘ ARKit ⓘ MetaHuman ⓘ 自定义

* 语音语种:

通用模式 中文语音 英文语音 日文语音

* 生成类型:

口型 口型+表情 (选择该项需要提交音频文件 + 表情标注文件)

表情动画渐入渐出设置 ⓘ : 动画开头 动画结尾

输出眼动控制动画曲线 ⓘ :

* 音频头部增加静音时间 ⓘ :

* 音频尾部增加静音时间 ⓘ :

确认提交

- 语音驱动 & 文本驱动 (提供多风格的TTS)
- 多语种: 中文, 英文, 日文等
- 7种默认表情 & 自定义表情动画曲线
- 细粒度配置: 眼动, 渐入渐出, 静音片段

Text2Face

驱动配置 ⓘ

* 音色:

爱小倩 爱小雪 爱小芸 爱小璐 爱小溪 爱小荷

爱小桃 爱小栋

* 帧率:

30fps 60fps

* 输出参数类型:

VISM16 ⓘ VISM16+AU20 ⓘ VISM16+ARKit ⓘ ARKit ⓘ MetaHuman ⓘ 自定义

* 生成类型:

口型 口型+表情 (请提交带表情标注的XML文件)

表情动画渐入渐出设置 ⓘ : 动画开头 动画结尾

输出眼动控制动画曲线 ⓘ :

确认提交

- 兼容自研及业界主流绑定标准
 - VISM16+AU20, ARKit BS, MetaHuman控制器
- 接入方案
 - WebUI, API, 离线调度, 流式服务

3D数字人AI驱动脸部动画 – 脸部动画曲线AI生成

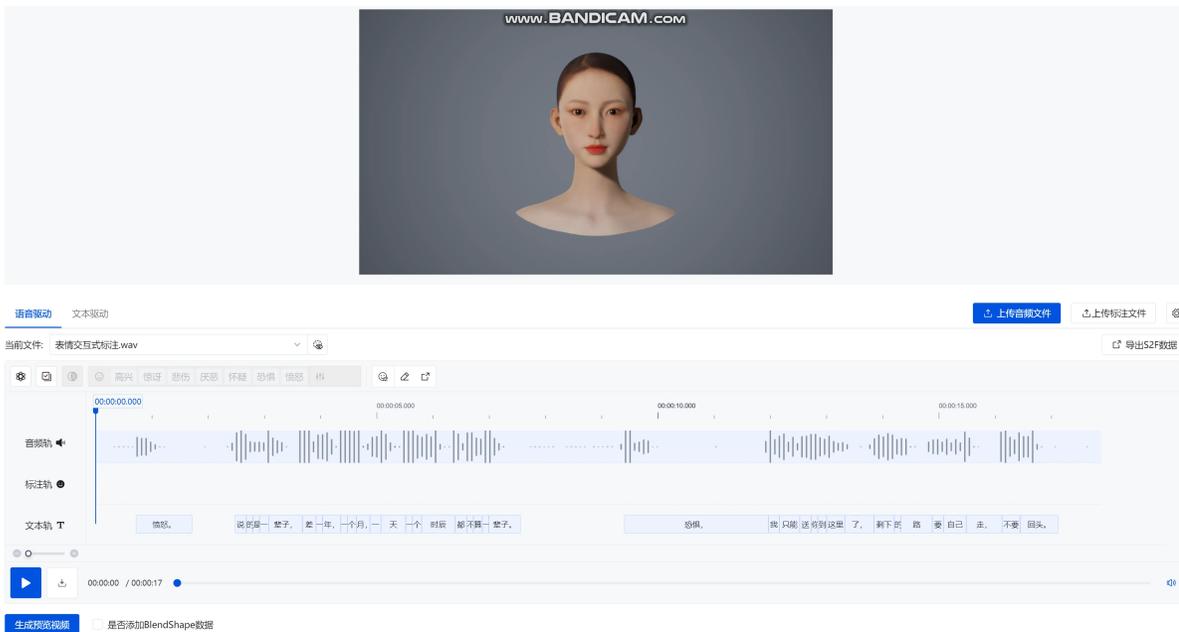
- 支持Speech2Face & Text2Face 驱动口型&表情
- 表情类型支持中立及7种表情，包括：
 - Angry (生气) Disgust (厌恶) Fear (恐惧) Happy (开心) Sad (悲伤) Surprise (惊讶) Suspicious (怀疑)
- Speech2Face支持中、英、日等多语种音频



3D数字人AI驱动脸部动画 – 交互式及智能表情标注工具

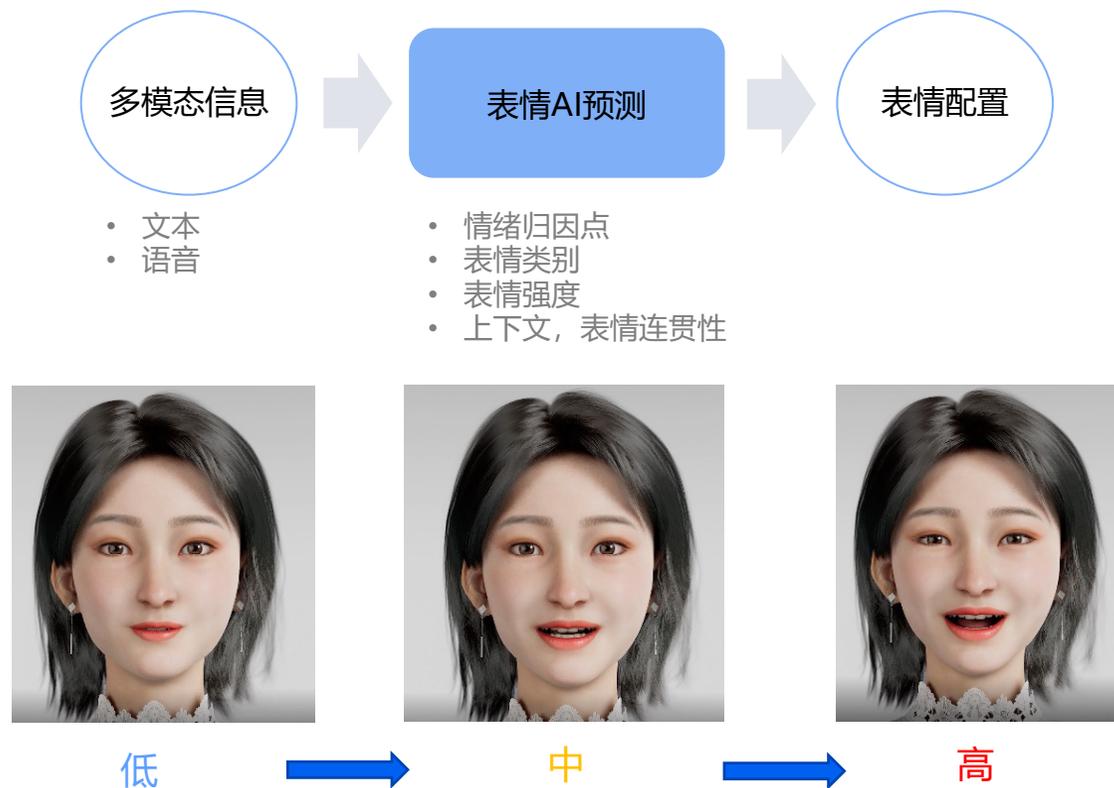
交互式标注

- 提供7种默认表情直接使用
 - Angry (生气), Disgust (厌恶), Fear (恐惧), Happy (开心), Sad (悲伤), Surprise (惊讶), Suspicious (怀疑)
- 支持自定义表情配置
- 提供全局快速标注及细粒度交互式标注
- 支持口型及表情的融合动态配置



智能标注

- 多模态 (语音信息, 文本信息)
- 多语种 (中文, 英文, 更多语种开发中)
- 默认表情即插即用, 自定义表情快速适配
- 面向角色、场景的智能标注自定义配置



3D数字人AI驱动脸部动画 – Pose及动画自定义编辑Maya工具

AutoRig 工具

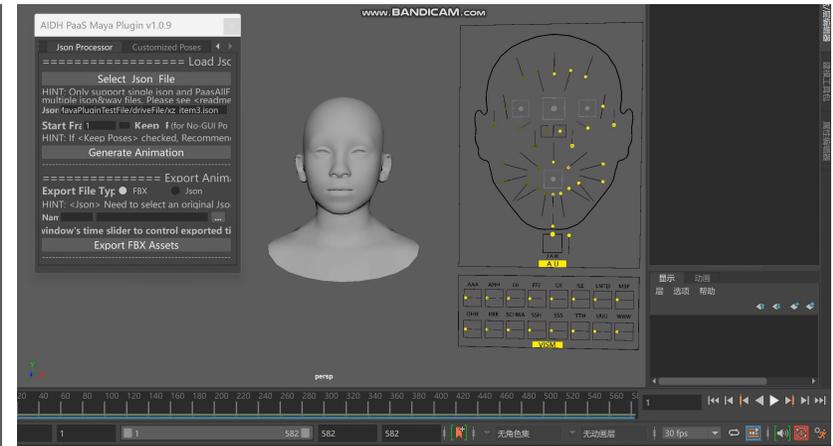
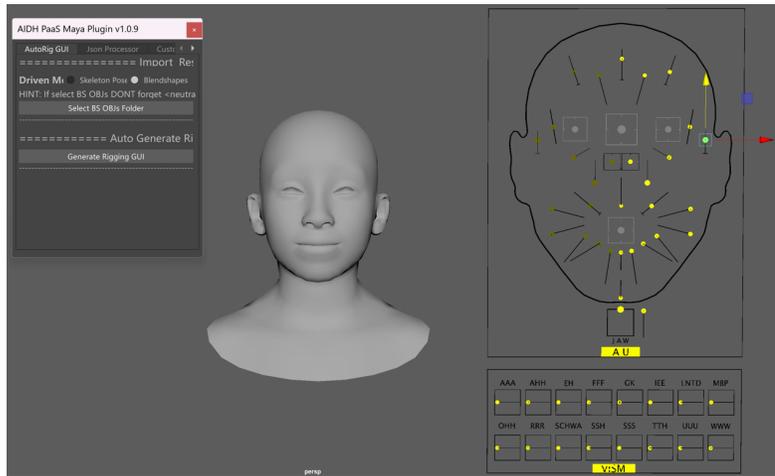
- 基于BS/骨骼Pose自动生成控制器面板
- 支持自研及业界主流绑定标准

自定义表情设置

- 可基于控制器设定自定义表情
- 自定义表情可配置于动画曲线AI生成能力，适配任意表情驱动

细粒度动画调整

- Human-in-the-loop工具，实现AI+人工的工具调试闭环，进一步提升动画的个性化表达需求及便捷度



1. 导入BS/骨骼Pose
2. 生成适配的控制面板

1. 定义自定义表情name
2. 激活控制面板，配置调整相应的表情表达

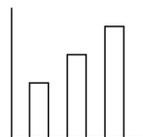
1. 导入AI生成动画曲线
2. 基于控制面板，per-frame细粒度调整动画，实现AI模式化表达（提效）+ 个性化表达（美学）的融合

3D数字人AI驱动脸部动画 – 业务价值

角色脸部动画生产方案对比



成本



效率



品质



自动化

全手K动画	高	低	高	否	品质上限, 不计成本, 难以scaleable
面捕+手K	中	低	中高	否	成本略低于手K, 管线冗长, 难以scaleable
口型及表情AI驱动	低	高	中高	可	兼顾效率与品质, 管线自动化、智能化

某3A游戏IP手游合作效能提升

降低角色动画制作**边际成本**

传统方案

~百/s

AI Flow

~0 /s

缩短角色动画资产**制作周期**

Manual Flow

角色脸部Pose

2 人天/角色

AI Flow

0.5 人天/角色

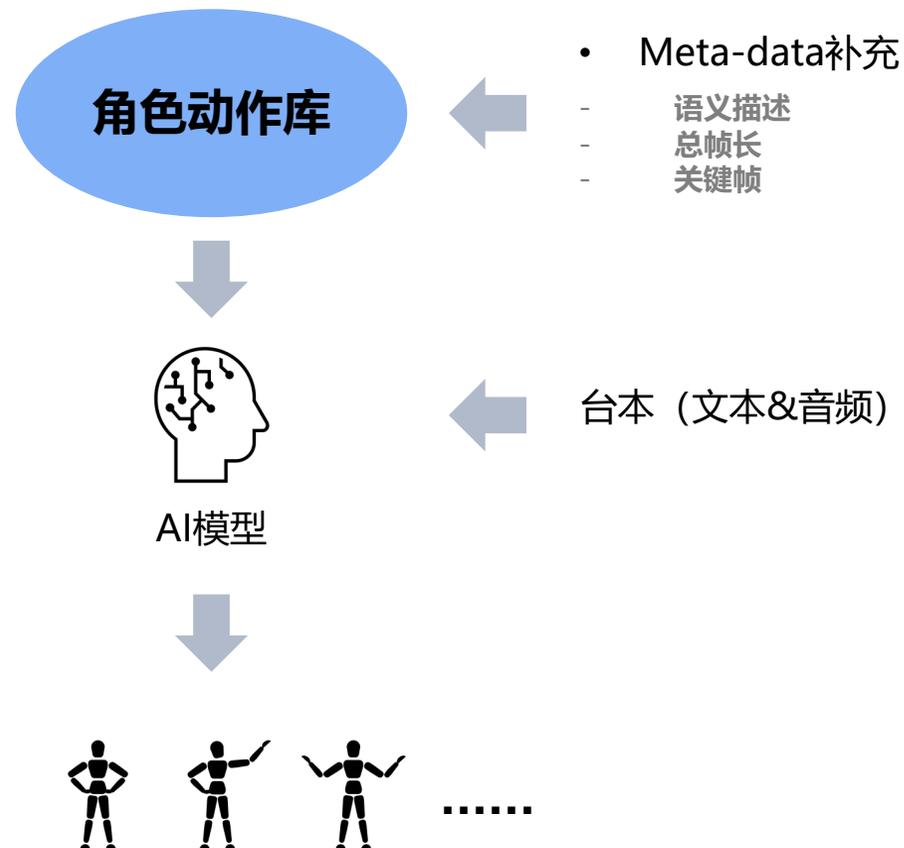
动画曲线

~0.5 人天/条

~几秒/段

3D数字人AI驱动身体动画 – AI动作召回&匹配

基于动作库的AI动作召回及匹配技术实现数字人动作驱动，是通过角色的语音及文本输入，通过AI模型去匹配并召回动作库中最适配的动作，再结合语音的节奏韵律和文本语义进行关键点对齐、速率和停顿控制，从而合成生动的虚拟人动作动画，提升3D数字人动画制作效率。



离线服务 vs 在线流式

离线服务

降低制作成本

提升工作效率

提高动画品质

智能化 管线化

在线服务



- NVIDIA ACE for Games Sparks Life Into Virtual Characters With Generative AI

端到端的数字人AI交互式驱动

语音/语义理解

LLM-based行为规划

语音合成

全身动作驱动

感谢倾听