



II 产品概述

云知声数字人通过音频和视频多模态融合、领先的人机对话交互全链路技术,结合语音识别、语义理解(大模型)、深度学习真人合成、有机融合视觉和听觉交互感受,提供高度模拟真人交互体验。广泛应用于交通客服导览、交通智能培训、勤务指挥监督、品牌形象代言等服务场景

II 产品介绍

+ 数字人形象

1. 真人形象1:1还原定制

先由真人拍摄视频,后通过AI算法训练生成,多形象、多音色可选,支持多套服装、站姿、坐姿,满足不同使用场景,录制3分钟视频完成形象定制,录制2分钟语音完成声音定制



支持多种生成方式 可采用云端生成视频和终端人机对话定制两种方式

2. 3D拟真/卡通数字人形象定制

高/低拟真的卡通数字人,需要先设计卡通形象然后进行3D建模和动作绑定,以精细/低拟的卡通形象与用户进行交互,整体制作周期一个月



3D真人形象定制 ▲



设计平面原画 确定形象外观概况

输出三视图 制作3D卡通模型

进行骨骼绑定 让模型"动"起来

将3D模型与外貌结合 渲染输出视频

3D卡通形象定制 ▲



支持多种虚拟形象

根据不同场景,应用卡通、真人等不同风格的虚拟形象可根据对话内容展现相应的表情、口型,如同与真人交互

3. 数字人引擎

口唇匹配

分析音频信号提取语音特征,实时驱动数字人口型模型变化

- 驱动口型运动
- 高度匹配语音输入
- 数百种口型模板
- 精准表达口型动作
- 真实流畅发音

面部驱动

利用深度学习技术,训练模型将面部数据映射至3D数字人模型,实现实时应用

- 表情自然逼真
- 语音内容驱动表情及肢体动作
- 丰富表情库和动作库
- 支持表情动作定制

动作触发

传感器捕捉惯性动作,AI实时渲染,语音、表情、情感与场景驱动,实现智能逼真3D数字人动作

- 动作触发驱动
- 动捕设备驱动
- 摄像头驱动
- 动作库

+ 感知交互

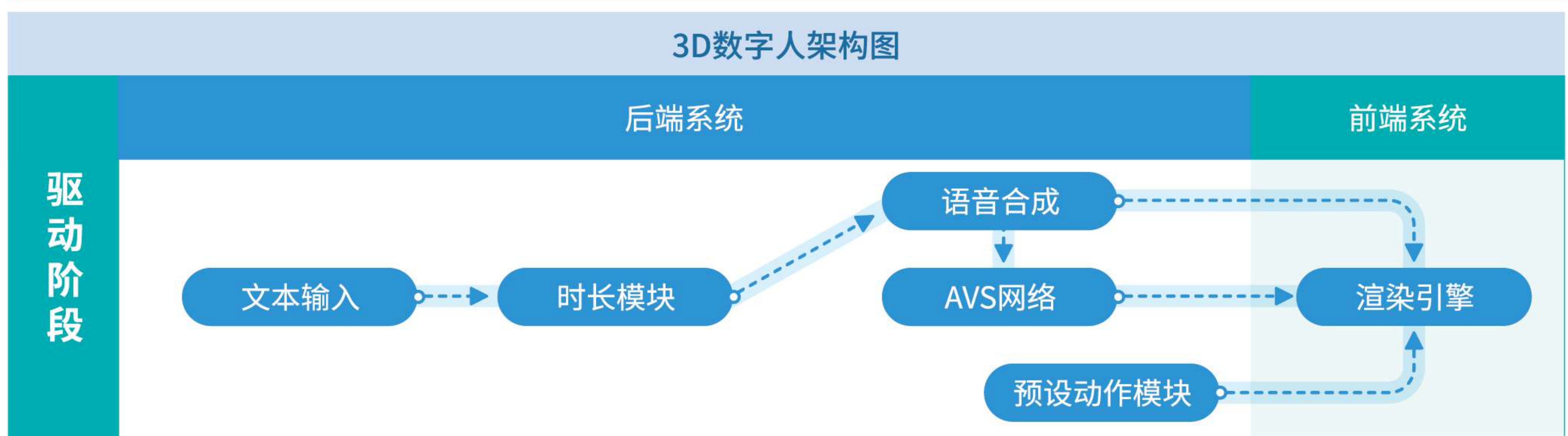


产品特性: 业界领先降噪算法确保音频纯净输入, 通过高精度ASR识别和大模型文字理解和意图识别能力, 最后通过端到端TTS技术将文字合成语音或复刻个性化声音, TTS在参数预测部分采用基于tacotron的seq2seg技术方案, 在声码器部分采用 WaveNet 的技术方案, 基于GAN和RNN 进行关键点预测和音视频生成

+ 认知交互

云知声山海大模型赋予数字人多模态融合认知大脑、感知声音融合、形成多模态交互闭环。应用到轨道交通、公共交通、航空交通、交通枢纽、道路交通等领域、实现知识检索查询、智能推理和比对、自动生成内容等智能化服务, 并配合插件API扩展功能与应用场景

+ 技术路线

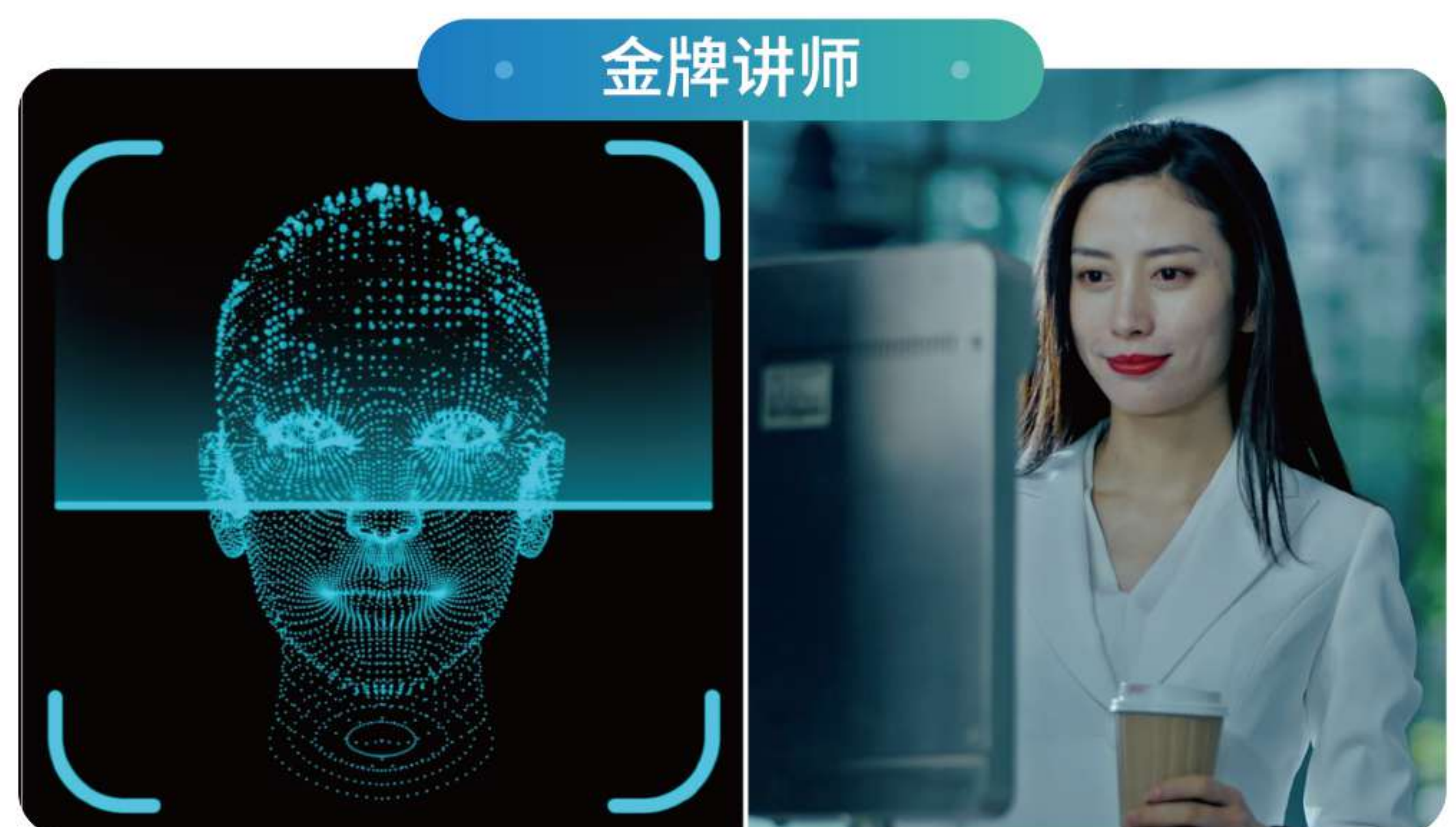


3D数字人技术路线 ▲

II 应用场景



提供服务咨询、站内引导、融合换乘指导、周边商家等功能



交通领域教育培训与模拟实训虚拟数字人方案



实现对勤务工作的全面监控、精准指挥和有效监督



青岛地铁吉祥物: 青小朵、青小通