



为什么随着技术发展，我们有了图文、语音、视频等远程交流方式，但在重要场合依然倾向于人与人之间面对面交流？不仅是因为直接交流舒服且自然，更重要的是面对面能够通过语气、态度，眼神、微表情、身体动作传递更多信息，更好更完善地表达自己的意思。

而这，也正是数字人划时代的地方，它既可以是虚拟世界的虚拟形象，也可以是人在元宇宙中的另一个分身，但无论是哪一种，都会模糊虚拟与现实的边界，为我们带来更为真实的交流互动体验。

因此，Rokid自研XR数字人技术的第一个核心是为用户提供自由“捏人”功能，其包含着完整的数字人设计系统。通过采用五头身比例，将写实风格与卡通风格进行平衡，让数字人的面部特征与表情动作更加突出和生动，同时也更具亲和力。再加上，五官和服装的自由搭配组合，可以让用户自由“捏”出更具个性的数字人形象。有消息称，Rokid还将推出更多的五官分类组合和动作，从而迎合全世界不同国家、不同肤色、不同文化的用户的审美。



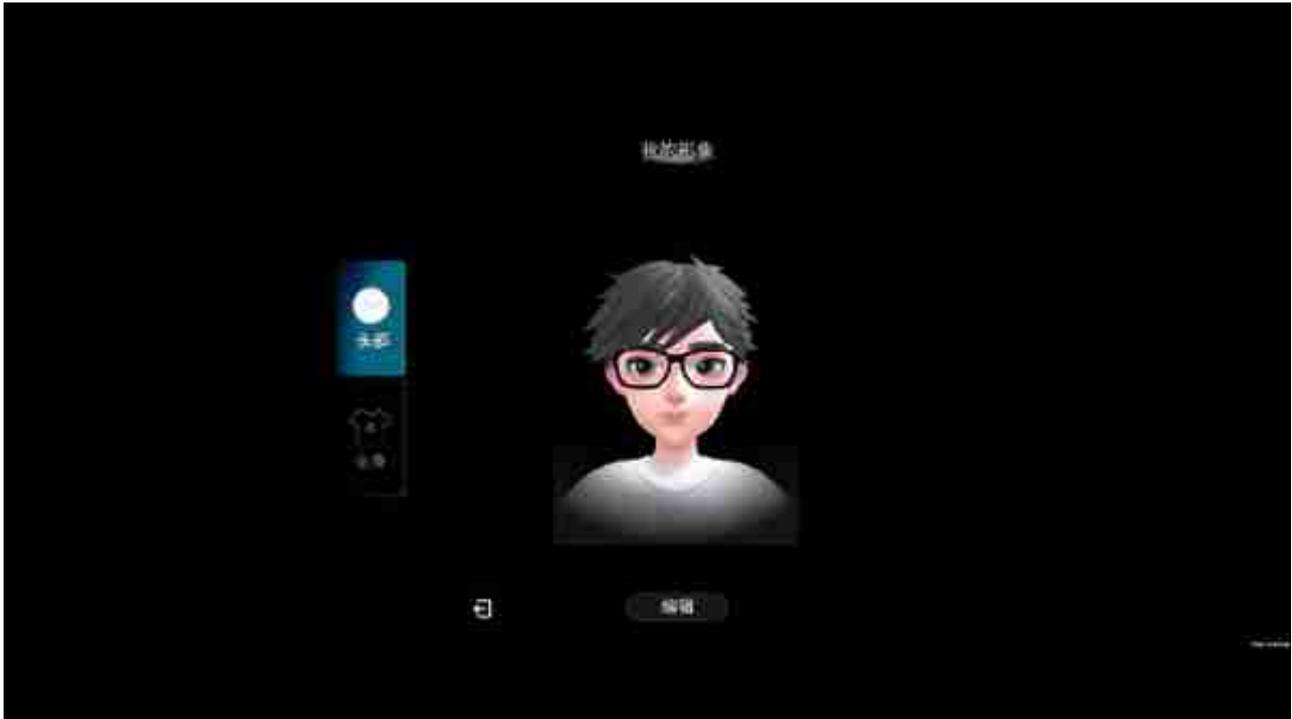
二、数字人产业的新支点

当前，业界实现数字人实时表情生成的主要方案，是面部捕捉与深度学习的结合。

这套方案的核心，是基于深度学习构建面部实时驱动系统，通过泛化的神经网络模块，根据真人表演录制的视频直接解算出3D角色的面部绑定动画参数值，从而实时驱动角色动画。这一方式较传统的面部捕捉方案更进一步，对设备的和算力的要求同样也会飙升，其成本可以达到百万元级别，而且还会随着建模精度等因素进一步提高，所以难以广泛普及应用。

而Rokid自研XR数字人技术，是根据音素——也就是根据语音的自然属性划分出来的最小语音单位，一个发音动作构成一个音素——来通过ASR（数字语音识别）算法倒推出发音时的口型动作，从而形成实时的表情动作，这样就摆脱了表情动作捕捉对摄像头等设备的依赖，实现口型与表情动作与真人同步匹配。

。



这一方案不仅降低了XR数字人的生成门槛，也为XR数字人产业带来了一次质变：它让数字人产业可以拓展到手机和AR设备上，为AR设备产业“注入灵魂”。

正如前文所述，XR数字人是连接元宇宙和用户的桥梁，只有当“桥梁”成本足够低，元宇宙相关应用和技术才能被用户广泛接受，并且成为推动AR等设备普及的“杀手级应用”和下一代互联网的支点。以Rokid发布的自研XR数字人技术为代表，我们正逐步接近一个技术“奇点”——数字人在C端低成本普及的大门正被缓缓打开，元宇宙的无穷魅力正在被更多人看到与参与。

三、让每个人都成为元宇宙的缔造者

回到这次技术发布，为什么Rokid自研XR数字人技术，着重强调了低成本与开放性？