



## 数字人克隆 - 视频采集规范

### 一、拍摄前准备

#### 1、场地要求

拍摄场地应选择安静、无噪声的环境，为保证拍摄器材架设方便，拍摄场地应相对宽敞。

#### 2、绿幕要求

- 为保证后期训练出的数字人能够自由替换背景，请准备绿幕进行拍摄，确保绿幕平整、颜色均匀无褶皱，可在开拍前使用挂烫机将绿幕熨烫平整。
- 使用绿幕背景拍摄视频时，若有拍摄全身的需求，地面也需要铺设绿幕。



图 1 数字人克隆场地示意图

#### 3、设备要求

- 选择一款可拍摄分辨率 1080P 及以上、帧率 25FPS 及以上的手机或摄像机；若使用手机拍摄，推荐直接用剪映录制；
- 准备固定手机或摄像机的三脚架，确保拍摄画面稳定、不失焦、无抖动；

- 准备补光灯（三盏最佳），保证均匀打光的光源，推荐使用“三点式打光”，为了保证抠图效果，可以准备一盏灯打轮廓光，要保证人物脸部，嘴巴，牙齿等五官细节清晰可见，无阴影；
- 准备提词器，需注意提词器架设位置，确保模特在读词时视线保持正视前方，视线不能偏移；
- 为了确保克隆出的数字人声音清晰、还原，需准备收声设备，如降噪麦克风、领夹麦等。



图 2 补光灯架设示意图

## 4、模特要求

### (1) 妆造要求

- 模特服装要求保持整洁，避免反光材质、半透丝纱材质、蕾丝材质的衣物；另外，服装颜色不要与绿幕相近，如黄色、绿色，以免影响抠图效果；
- 避免佩戴宽松易滑动晃动的首饰，比如项链、容易晃动的耳环（耳钉可以）等；如需佩戴眼镜尽量选择隐形眼镜或者小框镜，避免拍摄时镜面反光。
- 模特妆容干净、无忧光，拍摄前整理好发型，避免出现散发碎发以及发缝，



头发不要遮挡面部五官和脖颈，若有需要可使用发胶固定碎发。

## (2) 口播内容要求

- 在拍摄前，模特需提前熟悉口播文稿（5000字左右文稿），避免正式拍摄时频繁卡壳，文稿内容不限，需确保自己准备的文稿内容不重复，可以支持正常语速连续口播5分钟以上。

## (3) 动作要求

- 模特动作自然，动作尽量通用，避免具有语义或指向性的动作，需能够适配所有场景；
- 避免与讲话无关的表情，例如皱眉、挑眉、撅嘴、舔嘴、吞咽等；
- 注意眼神直视前方，不乱瞟，不张望，避免抬头或低头注视提词器；
- 避免身体晃动，不要做大幅度和迅速的动作，手不挡脸，抬手不要过肩。

## (4) 其他要求

- 模特站立需距离绿幕2米以上，避免绿幕颜色浸染，影响抠图效果；



图 3 拍摄模特示意图

## 二、拍摄过程

### 1、试拍测试

- 正式拍摄前，调试好灯光及拍摄设备，可让模特进入拍摄区，确保光线均匀、无阴影、不过曝；
- 模特在镜头画面中比例要合适，可引导模特进行调整，确保目光视前方，不俯仰、不偏斜，手部做动作时不出画面；
- 模特根据提词器试读文稿，调整提词器位置及滚动速度，确保模特能够以正常语速顺畅阅读；
- 确保拍摄现场安静、无噪音，录制的视频模特声音清晰、音画同步；
- 若模特欠缺拍摄经验，可引导其自然动作，例如双手自然下垂或者交握于身前，确保模特状态正式、自然、不僵硬。



## 2、正式拍摄

- 正式拍摄需清场，一个训练视频拍摄时间约 5-6 分钟，拍摄需一镜到底，中间不允许暂停或后期剪辑；
- 拍摄前十秒及后十秒需保持静止不动，手可放置于身前，保持直视镜头，不要张嘴或做其他动作；
- 开始五分钟口播后，启动提词器滚动播报文本，模特需保持平视前方录制口播视频，在阅读文稿的过程中，需保持吐字清晰，语速不要过快或过慢，句与句之间可以有 1-2 秒间隔，断句清楚。

**注意：内容读错不要紧，继续往下读即可，不要重复说错的内容。**

## 三、训练视频提交

- 数字人训练视频拍摄完成后，最终需要提交的是一个带音轨的视频文件，**视频要求尺寸 1080\*1920px，mp4 格式，25 帧率，时长 5-6 分钟。**
- 训练视频提交至 AI 数字人元启平台，等待审核、克隆完成即可。



图 4 训练视频上传平台示意图

## 四、数字人形象定制须知

- 数字人最终呈现效果与您上传的数字人训练视频中模特的形象、动作基本一



致，请确保上传的训练视频任务形象为个人最理想的状态；

- 视频提交后，我们会对您的视频进行审核，若您提交的视频不符合定制要求，我们将告知您审核失败的原因，并需要您按照拍摄规范重新拍摄、上传视频。
- 您提交的视频审核通过后，即可进入数字人克隆流程，数字人训练完成大概需要 **1-2 个工作日**，请您耐心等待；
- 数字人形象克隆完成后，您将获得该形象的使用权，我们承诺将对您上传的视频进行隐私处理。