**简介**

基于计算机视觉和自然语言处理的实时跟拍、在线直播平台。

* **详情描述**

通过人类自然语言的指导，指导摄像机器自动追踪，结合识别，定位，跟踪，调焦等摄影技术，通过文字声音图片录像全方位多角度地展现赛场或者会场的精彩情况

通过解说员的解说词，通过语音识别和自然语言处理技术来确定对那个物体进行特写，并利用计算机视觉技术和硬件控制来**定位识别追踪**相关选手，进行**自动跟踪**拍摄，以求减少摄影师及后期剪辑的人力物力支出，提高效率。

* **创意来源**

平日里看球赛、新闻、综艺节目的直播，现阶段我们通常采用的技术都是：主持人发表评论或看法，现场的多个摄影师分工跟踪拍摄特定的人物，导播室的工作人员结合主持人的解说，经过后期的人工处理，挑选切换播报给观众的镜头至合适的视频画面。这样的问题在于，直播这样一场比赛，需要大量的人力物力，另外，解说要迁就摄像导播，不能尽情的控制整个解说的流程。但有了这样一套系统之后，费用将大大降低，通过文字图片，以及更加贴近运动员的录像，从全方位解读比赛，使得直播系统在日常训练赛，和学生比赛中更显神威

今天在HACKTHON X FDU活动开幕的现场，看着扛着长枪短炮各种给大家摄影的辛苦的摄影师，我们突然想到，这件事可不可以让人工智能来帮我们完成呢？

* **解决的问题**

对重大赛事减少在媒体转播上面所支出的人力物力，并实时处理多媒体信息

基本实现系统全自动，除重大赛事外，也能用于训练赛，为教练提供宝贵的训练资料，用于学生比赛，记录青春的精彩瞬间。用于召开的会议，不放过大会的每一个关键点

可以成为比鹰眼更加全方位的裁判系统

在网络教育发达的新时代，为网课制作和远程公开课提供了技术支持

* **项目使用指南**
  + **预先训练**
  + 使用步骤

**功能设计说明**

**视觉交互设计说明**

运用计算机视觉上的级联分类器和物体追踪技术，对视频流进行处理

* **核心技术说明**

语音识别及自然语言处理

**1，自然语言处理，使用python编写录音程序，并定时形成录音文件。使用IBM提供的Watson接口进行处理。**

计算机视觉

用Matlab CV toolkit 进行人脸及其他物体的识别框选追踪定位

硬件交互

* + 硬件部分由Arduino UNO 开发板、罗技C920摄像头、小车组件构成。
  + Arduino由12VDC供电，负责与MATLAB作串口通信，以9600波特率传输响应信号。
  + 摄像头实时采集视频图像，通过USB3.0传输到MATLAB分析。
  + 因此Arduino开发板与小车和MATLAB构成闭环信号系统。
  + 当MATLAB分析图像得出目标物体处于偏左/右结果时，通过USB发出左/右转信号，信号到达开发板后，输出到9g舵机，带动摄像头整体向左/右转。
  + 小车组件同时由Arduino开发板驱动，可以向前，向后，左转，右转。
* **SDK使用情况说明**
  + 语音识别：
    - IBM WATSON
  + 自然语言处理
    - IBM WATSON
  + 计算机视觉
    - Matlab CV tookit
* **未来发展规划**

对于赛事

完善整一套软硬件直播系统，小车的数量增加，甚至可以引入无人机，多角度，多方位记录比赛。计算视觉的物理识别追踪的硬件集成在终端，识别的物体也更加多样，如可以球衣颜色和号码来识别特定的运动员和追踪。在识别定位精度上面，可以通过添加全局远焦摄像机，来粗调选手位置，在终端继续定位追踪细调选手位置，做到远近有致，可捕捉运动员的慢动作回放也可捕捉观众的喝彩瞬间。解说语音所转成的文字和自动摄像所拍下的精彩瞬间可以轻松做成图文直播

* 对于网课录制和远程教育，

当老师说处关键词比如“大家注意”，”这里难懂“的时候，我们的镜头自动放大，同时跟踪老师，上课回答问题的同学可以给特写 ，捕捉互动的精彩瞬间。 此套系统可以录制网课，也可进行实时远程公开课。

* **相关文件**
* **DEMO展示地址**
* **代码仓库地址**
* **视频展示**
* **图片展示**
* **PPT展示**