**Vision详细设计文档**

目录

[1 硬件采集模块 3](#_Toc104037155)

[1.1 功能描述 3](#_Toc104037156)

[1.2 性能描述 3](#_Toc104037157)

[1.3 输入 3](#_Toc104037158)

[1.4 输出 3](#_Toc104037159)

[1.5 程序逻辑 4](#_Toc104037160)

[1.6 限制条件 4](#_Toc104037161)

[2 设备发现连接模块 4](#_Toc104037162)

[2.1 功能描述 4](#_Toc104037163)

[2.2 性能描述 5](#_Toc104037164)

[2.3 输入 5](#_Toc104037165)

[2.4 输出 5](#_Toc104037166)

[2.5 程序逻辑 5](#_Toc104037167)

[2.6 限制条件 6](#_Toc104037168)

[3 视频模块 6](#_Toc104037169)

[3.1 功能描述 6](#_Toc104037170)

[3.2 性能描述 6](#_Toc104037171)

[3.3 输入 6](#_Toc104037172)

[3.4 输出 6](#_Toc104037173)

[3.5 程序逻辑 7](#_Toc104037174)

[3.6 限制条件 7](#_Toc104037175)

[4 视频融合模块 7](#_Toc104037176)

[4.1 功能描述 7](#_Toc104037177)

[4.2 性能描述 8](#_Toc104037178)

[4.3 输入 8](#_Toc104037179)

[4.4 输出 8](#_Toc104037180)

[4.5 程序逻辑 8](#_Toc104037181)

[4.6 限制条件 9](#_Toc104037182)

[5 人脸识别模块 9](#_Toc104037183)

[5.1 功能描述 9](#_Toc104037184)

[5.2 性能描述 9](#_Toc104037185)

[5.3 输入 9](#_Toc104037186)

[5.4 输出 9](#_Toc104037187)

[5.5 程序逻辑 10](#_Toc104037188)

[5.6 限制条件 10](#_Toc104037189)

[6 图片拼接模块 10](#_Toc104037190)

[6.1 功能描述 10](#_Toc104037191)

[6.2 性能描述 11](#_Toc104037192)

[6.3 输入 11](#_Toc104037193)

[6.4 输出 11](#_Toc104037194)

[6.5 程序逻辑 11](#_Toc104037195)

[6.6 限制条件 12](#_Toc104037196)

[7 全景融合模块 12](#_Toc104037197)

[7.1 功能描述 12](#_Toc104037198)

[7.2 性能描述 12](#_Toc104037199)

[7.3 输入 12](#_Toc104037200)

[7.4 输出 12](#_Toc104037201)

[7.5 程序逻辑 12](#_Toc104037202)

[7.6 限制条件 13](#_Toc104037203)

[8 蓝牙连接模块 13](#_Toc104037204)

[8.1 功能描述 13](#_Toc104037205)

[8.2 性能描述 13](#_Toc104037206)

[8.3 输入 14](#_Toc104037207)

[8.4 输出 14](#_Toc104037208)

[8.5 程序逻辑 14](#_Toc104037209)

[8.6 限制条件 14](#_Toc104037210)

[9 舵机控制模块 15](#_Toc104037211)

[9.1 功能描述 15](#_Toc104037212)

[9.2 性能描述 15](#_Toc104037213)

[9.3 输入 15](#_Toc104037214)

[9.4 输出 15](#_Toc104037215)

[9.5 程序逻辑 15](#_Toc104037216)

[9.6 限制条件 18](#_Toc104037217)

# 硬件采集模块

## 功能描述

此模块实现图像数据的采集以及与传输模块的对接功能。具体包括高清拍照后台服务的设计、高清录像存储、转码设计、采集与传输机制的对接。

高清拍照后台服务独立于推流服务，实现与Android底层摄像机接口的使用，为用户实现高清相片体验（实际分辨率取决于相机硬件，已测试可实现4K分辨率）。控制端对采集端发送拍照指令，采集端与控制端建立文件传输通道之后进行传送。

高清录像存储集成自推流采集端，可满足1080P的高清视频数据采集以及时间截取。控制端对采集端传送开始和结束指令，接到结束指令后，采集端向控制端建立文件传输通道之后进行传送。

## 性能描述

拍照分辨率可达4K，录像分辨率可达1080P。

## 输入

控制端：界面按钮响应、采集端建立请求与图像传输流

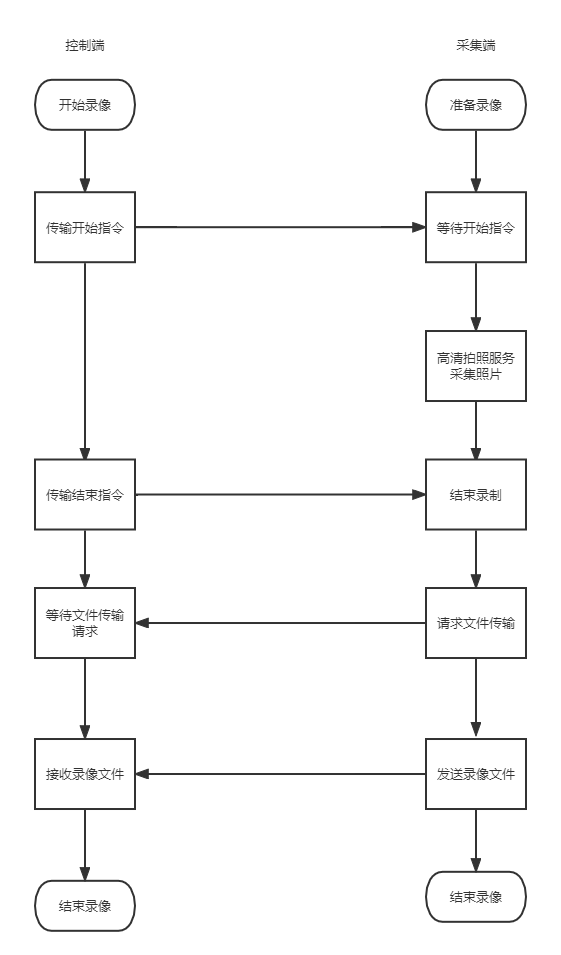
采集端：界面按钮响应、控制端指令

## 输出

控制端：控制端指令、图像结果输出

采集端：采集端图像结果、图像传输请求与图像传输流

## 程序逻辑



## 限制条件

相机硬件适配、相机服务稳定

# 设备发现连接模块

## 功能描述

此模块利用WiFiP2p技术实现设备之间的发现与互连。连接成功后可以实现指令和文件的传输。

我们的控制端设备将作为服务器，首先作为群主创建群组，然后创建ServerSocket等待客户端连接，连接到客户端Socket后持有客户端的IO流，就可以实时传输指令和文件了。

我们的采集端将作为客户端，首先初始化模块，然后注册广播，异步扫描范围内的群主设备，通过广播异步获取扫描到的设备列表。扫描完成后，选择设备进行连接，构建Device config异步连接某个设备，通过广播接收连接状态信息，连接成功后，将设备信息返回服务端。同时连接成功后可以获取服务端IP地址，从而可以与服务端建立Socket通信，实现指令和文件的传输。

## 性能描述

同一个局域网下，扫描延迟在500ms内，连接延迟在800ms内

## 输入

扫描：无

连接：要连接的设备Config

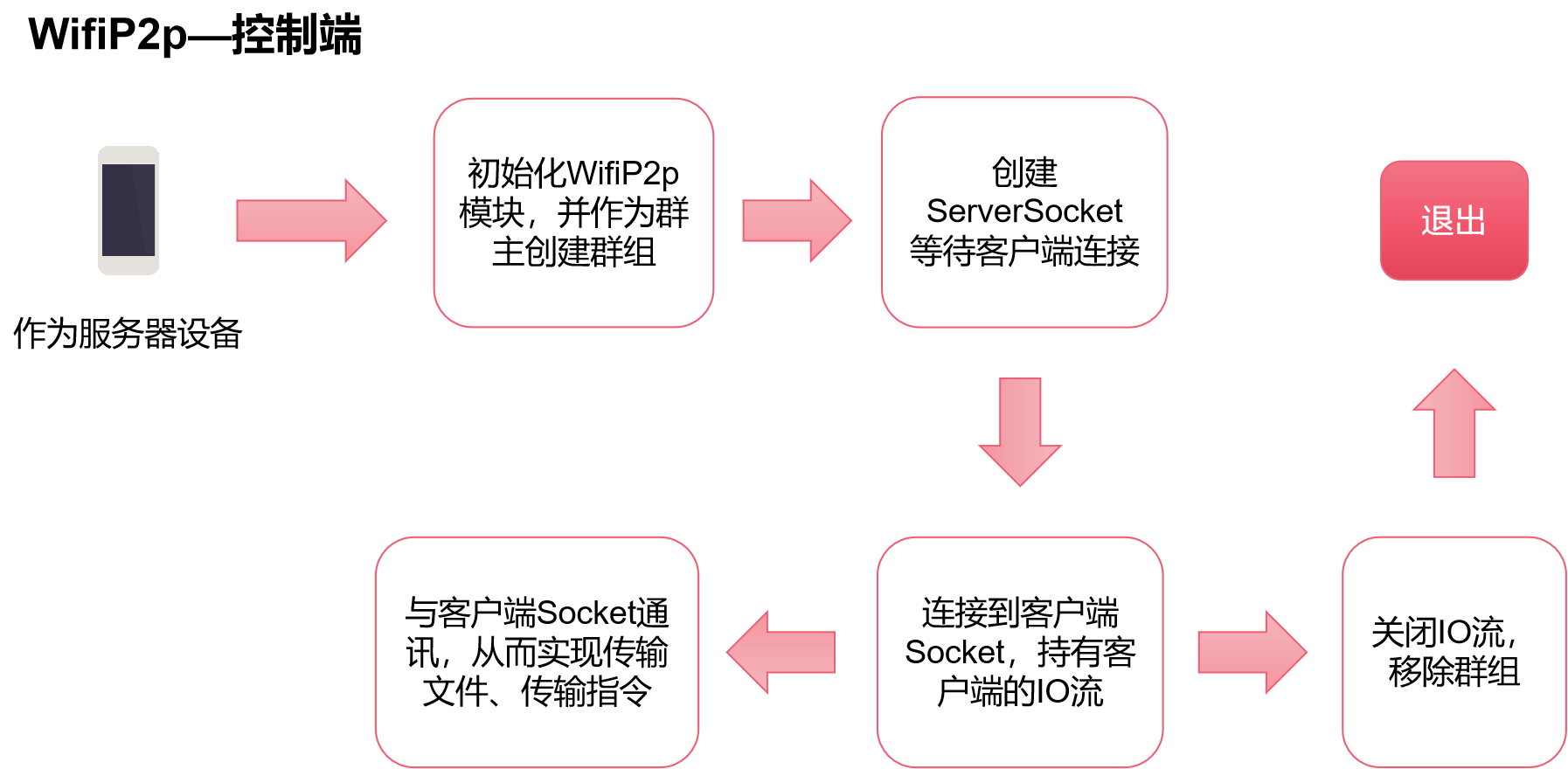
## 输出

扫描：扫描到的设备List

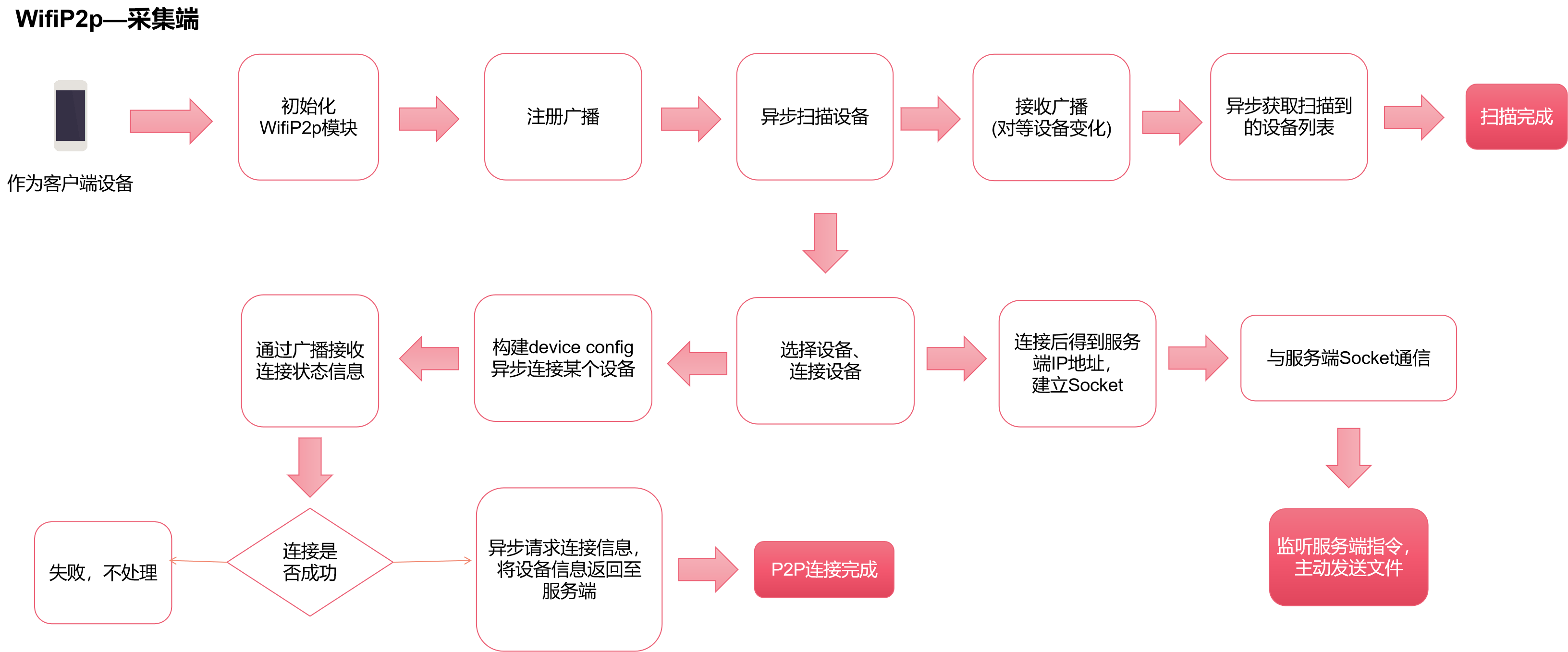
连接：连接成功后输出连接设备信息，失败输出失败原因

## 程序逻辑

控制端



采集端



## 限制条件

设备支持WifiP2p，设备打开wifi并连接到同一局域网

# 视频模块

## 功能描述

控制端和采集端在确定连接之后，本模块提供设备间视频信息的交换功能，及所谓的推流，并保证一定的视频质量和较短的时延。

我们首先需要获取本地的视频流信息，紧接着需要建立连接通道，WebRTC是基于P2P的，但是在连接通道建立好之前，我们仍然需要服务器帮助传递信令，而且需要服务器帮助进行网络穿透，我们通过实现了ICE框架的STUN服务器进行内网穿透，使设备间可以进行P2P视频传输，随后便是将视频流通过点对点进行传输。

## 性能描述

该功能的视频质量可达720p，时延控制在800ms内

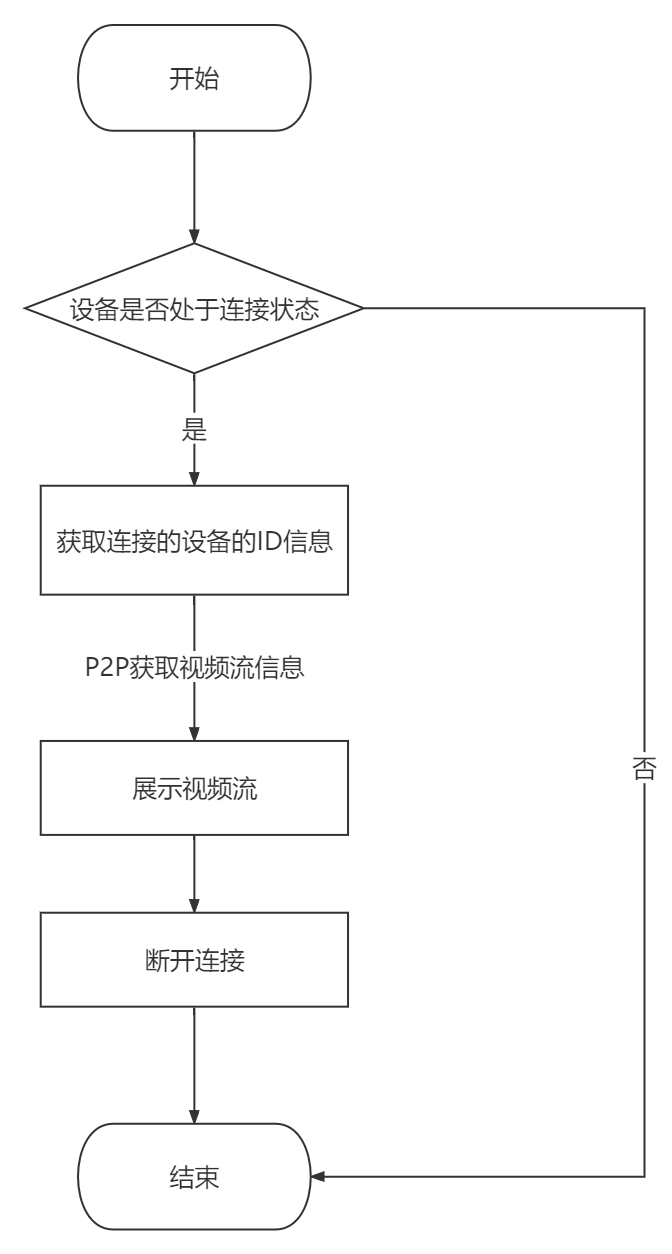
## 输入

设备Id号 userId：String

## 输出

视频流数据 stream：MediaStream

## 程序逻辑



## 限制条件

设备间处于连接状态

# 视频融合模块

## 功能描述

本模块在控制端获取到采集端视频数据后，通过ffmpeg技术对视频和图片信息进行处理，实现视频融合的功能。

视频融合部分，当我们有一组拍摄同一事物的录像需要融合时，我们需要确定每个时间点在录制的视频文件，因此在录像时，切换录制对象的时候，我们记录下上一个片段的开始时间和持续时间并存储下俩，并将开始时间重新设置为开始时间+持续时间，持续时间设置为0，用于下一次截取。当获得了片段时间点信息后，我们便可以使用ffmpeg对视频进行融合操作。

## 性能描述

融合视频时间不超过3s

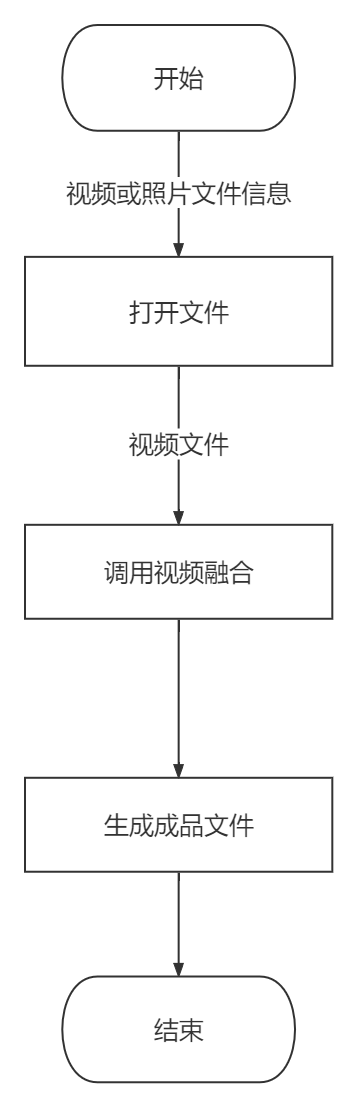
## 输入

视频文件 File

## 输出

融合视频文件 File

## 程序逻辑



## 限制条件

无

# 人脸识别模块

## 功能描述

进入采集端视频显示页面时，人脸识别模块会自动启动，通过opencv寻找拍摄图像上的人脸信息，并返回人脸的坐标位置，用于后续蓝牙连接模块对采集端设备位置的调整。

## 性能描述

每帧图片处理时间不超过50ms

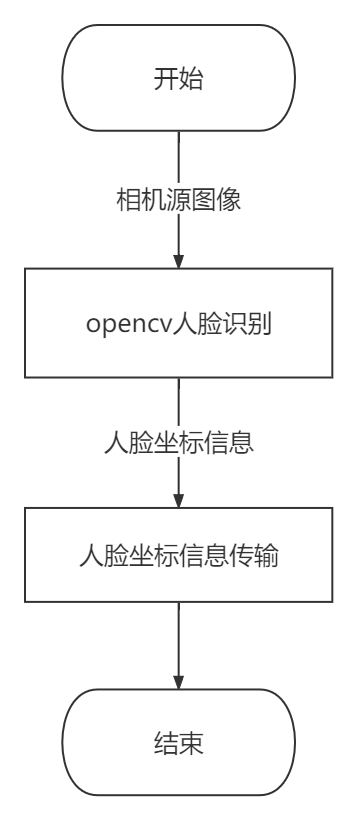
## 输入

相机捕捉到的源图片

## 输出

人脸出现的坐标位置

## 程序逻辑



## 限制条件

相机权限开启。

opencv模块加载。

读写手机内存权限开启。

# 图片拼接模块

## 功能描述

进入拼接图片模式后，用户点击拍照，程序会自动调用拼接图片方法，并将拼图后的结果直接存储到本地相册，用户可实时查看。

## 性能描述

拼图处理和存储结果到相册时间总计不超过100ms

## 输入

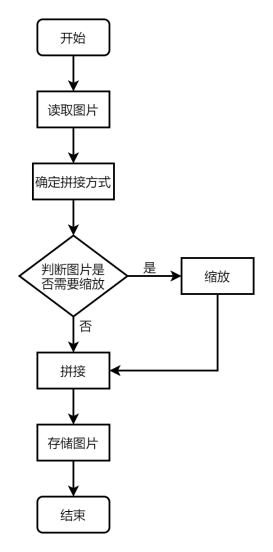
需进行拼接的图片路径和名字

## 输出

输入的所有图片成功拼接后的图片

## 程序逻辑

用户点击拍照后，程序根据存储到本地的文件路径和名字获取需要拼接的图片，根据图片数量确定拼接方式，并计算宽高比例对相应图片进行缩放，然后利用Canvas绘制bitmap进行拼接，最后存储拼图结果到本地相册。



## 限制条件

模式为拼接图片模式；

读写手机内存权限开启。

# 全景融合模块

## 功能描述

进入全景融合模式后，用户点击拍照，程序会自动调用全景融合方法，并将融合后的结果直接存储到本地相册，用户可实时查看；若图片特征点不足导致无法融合会返回提示信息。

## 性能描述

全景融合和存储结果到本地的时间总计不超过900ms

## 输入

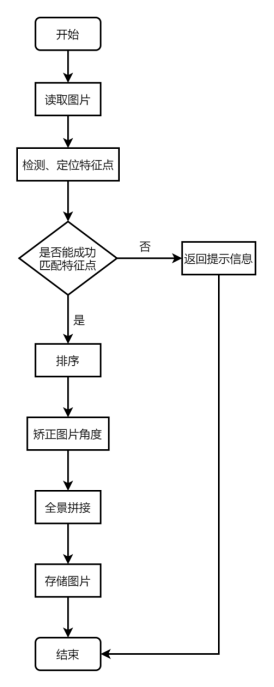
需进行全景融合的图片路径

## 输出

输入的图片成功融合为全景后的图片或提示信息

## 程序逻辑

进入全景融合模式后，用户点击拍照，程序根据文件路径获取需要融合的图片，利用OpenCV Stitcher检测和定位图片特征点，对图像特征点进行匹配，并对图像进行排序，矫正角度后进行全景拼接，最后存储拼接后的全景图到本地相册。若特征点匹配度不足，将返回提示信息。



## 限制条件

模式为全景融合模式；

读写手机内存权限开启。

# 蓝牙连接模块

## 功能描述

在采集端进入采集画面前，通过蓝牙连接舵机控制板，进而能完成后续对舵机的控制操作

## 性能描述

蓝牙连接时间小于5s

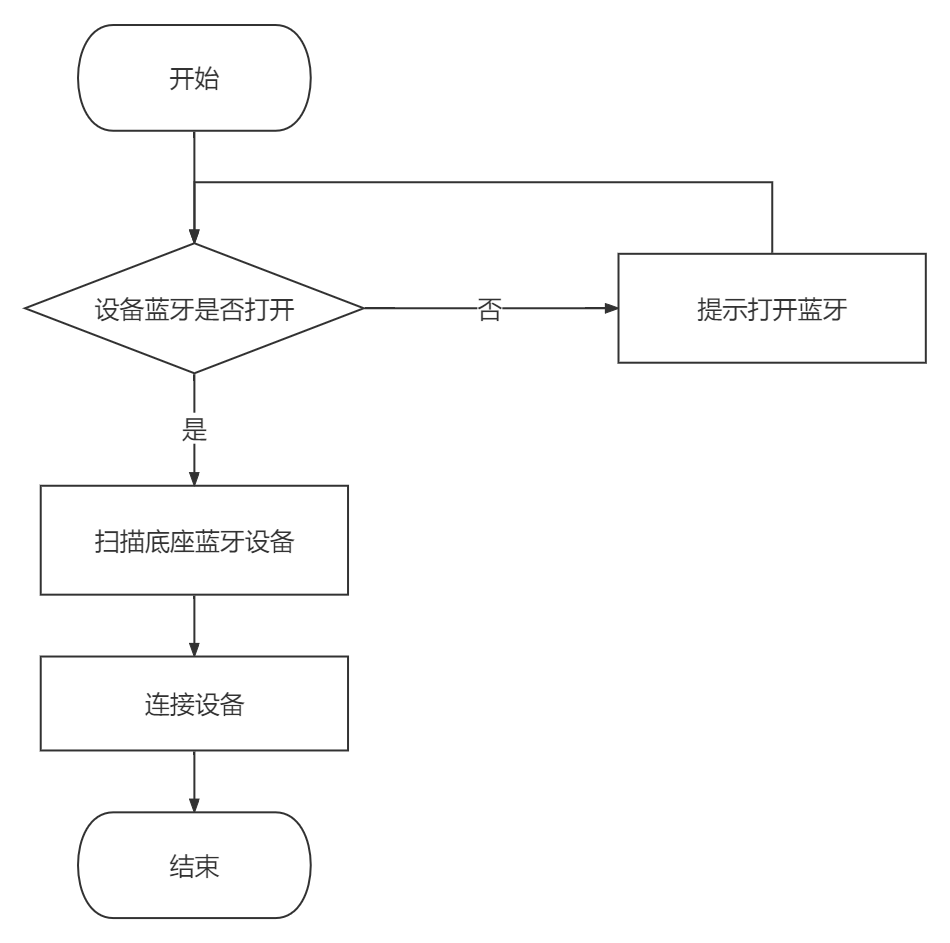
## 输入

舵机控制板的蓝牙设备名

## 输出

连接成功信息

## 程序逻辑



## 限制条件

手机蓝牙功能启动

手机舵机蓝牙功能可用

# 舵机控制模块

## 功能描述

舵机控制模块在获取人脸位置信息之后，通过PID算法确认舵机的角度变化需求，然后根据串口协议编写并发送指令到舵机，使舵机完成旋转功能。

## 性能描述

发出请求到舵机完成旋转时延小于500ms

## 输入

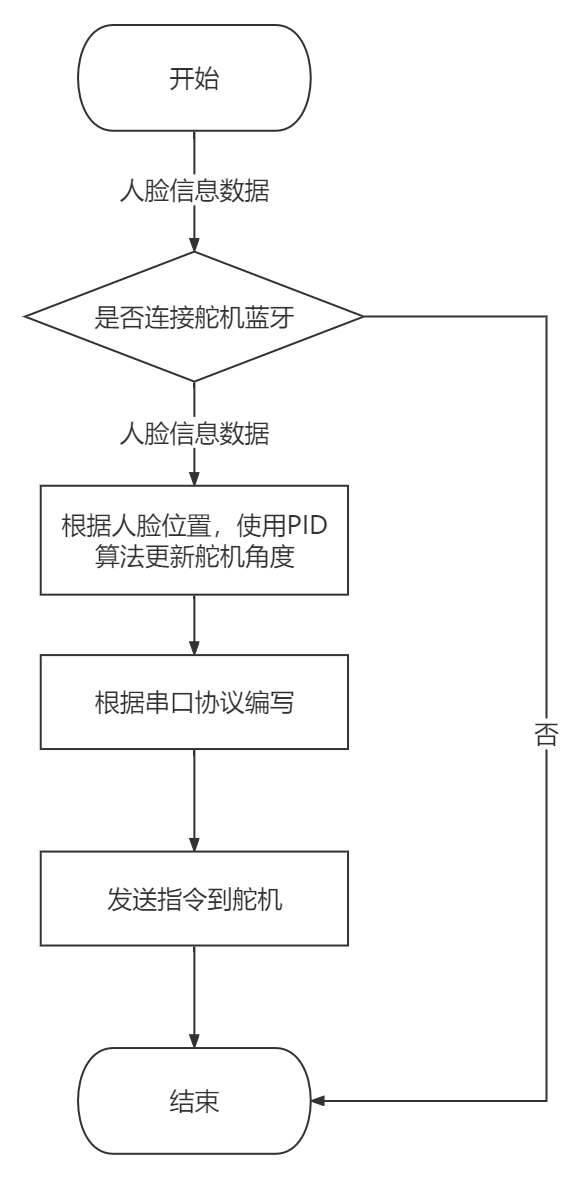
人脸位置信息

## 输出

编码指令及舵机旋转变化

## 程序逻辑

1. 整体流程

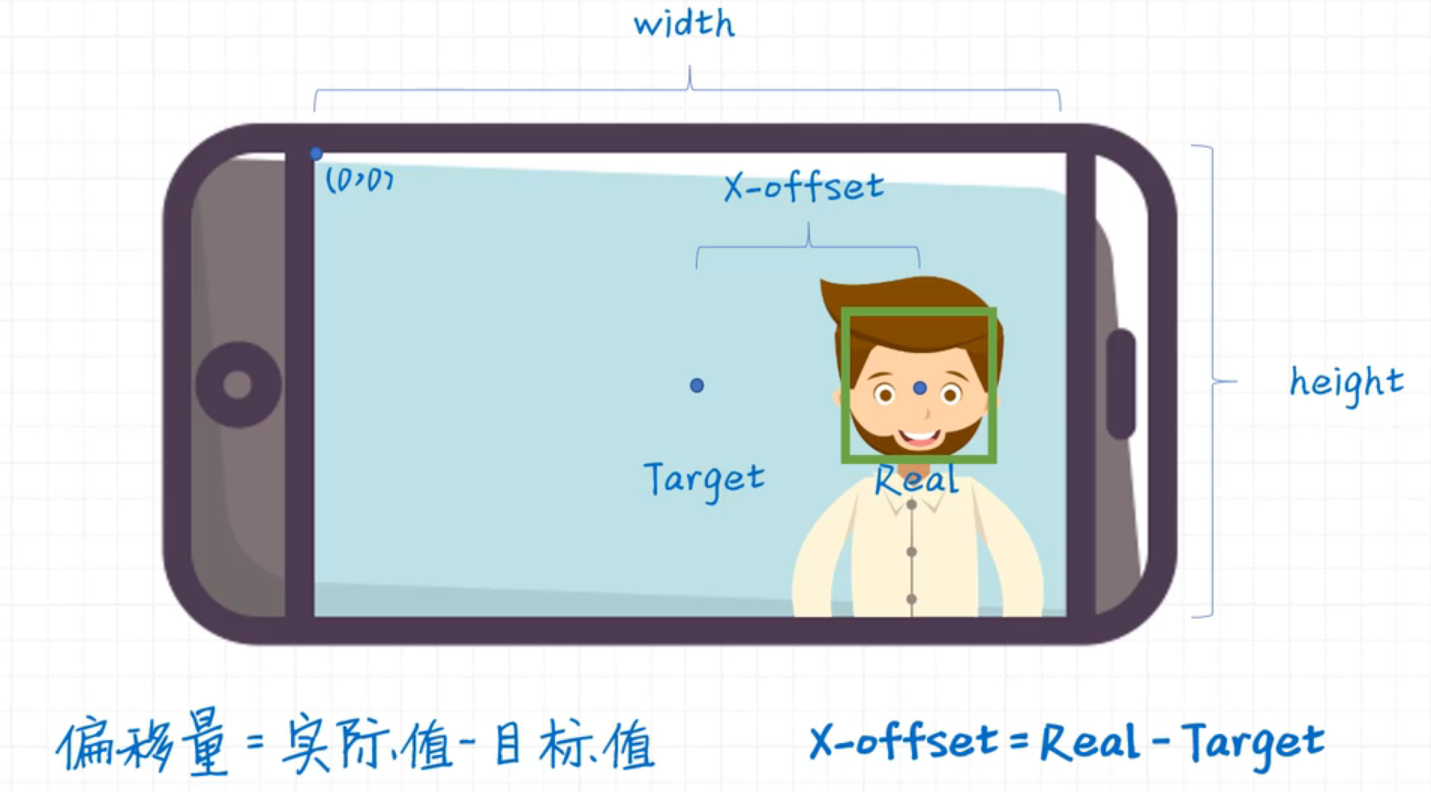
.

2. PID控制

PID控制算法是自控里最经典，使用最广泛的算法，PID的全称为比例-积分-微分控制器（Proportion-Integration-Differentiation）。

这里我们只用到了比例控制。

1. 像素偏移量

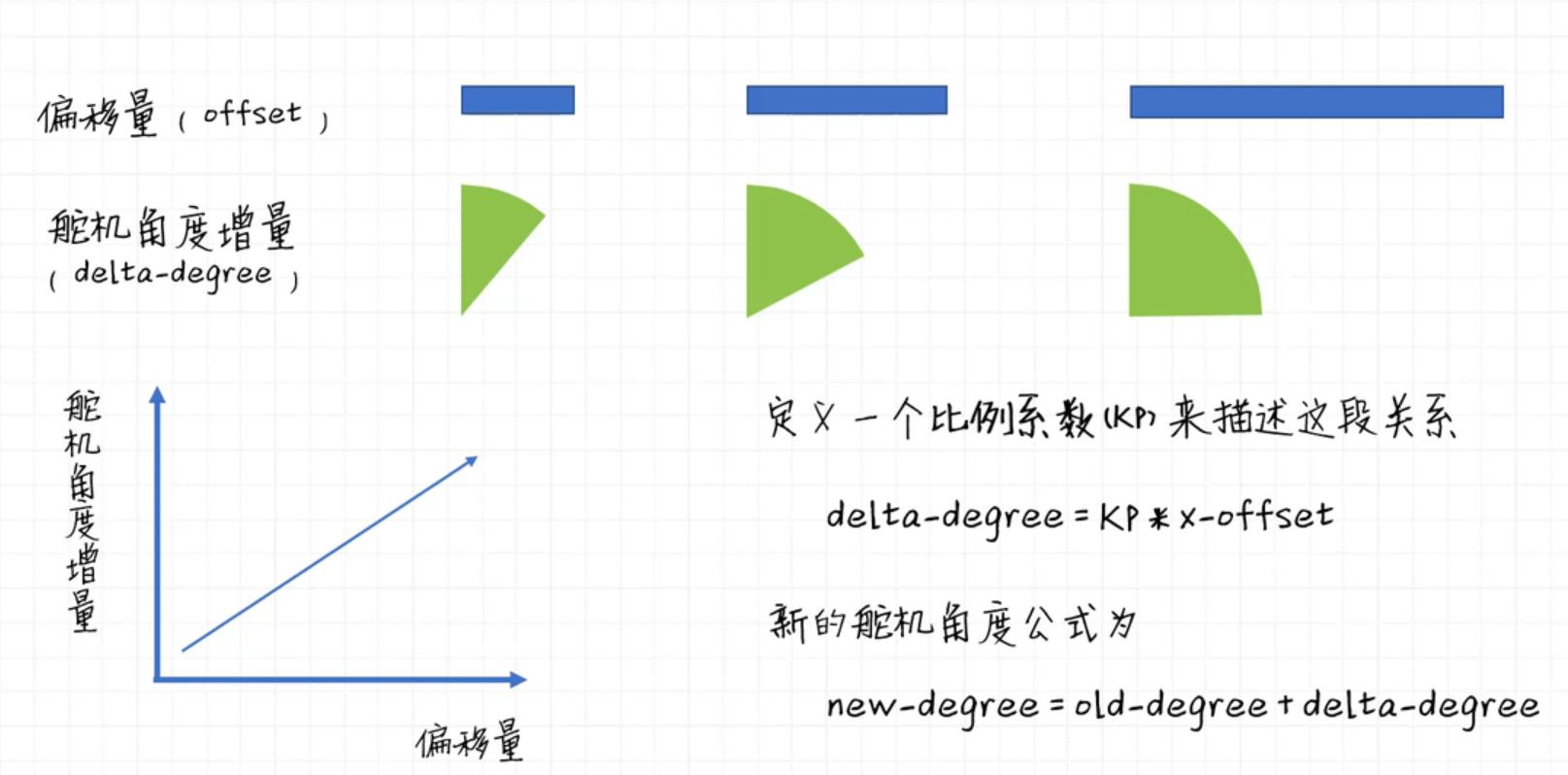


根据获取到的box值计算像素偏移量X-offset

1. 偏移量归一化处理

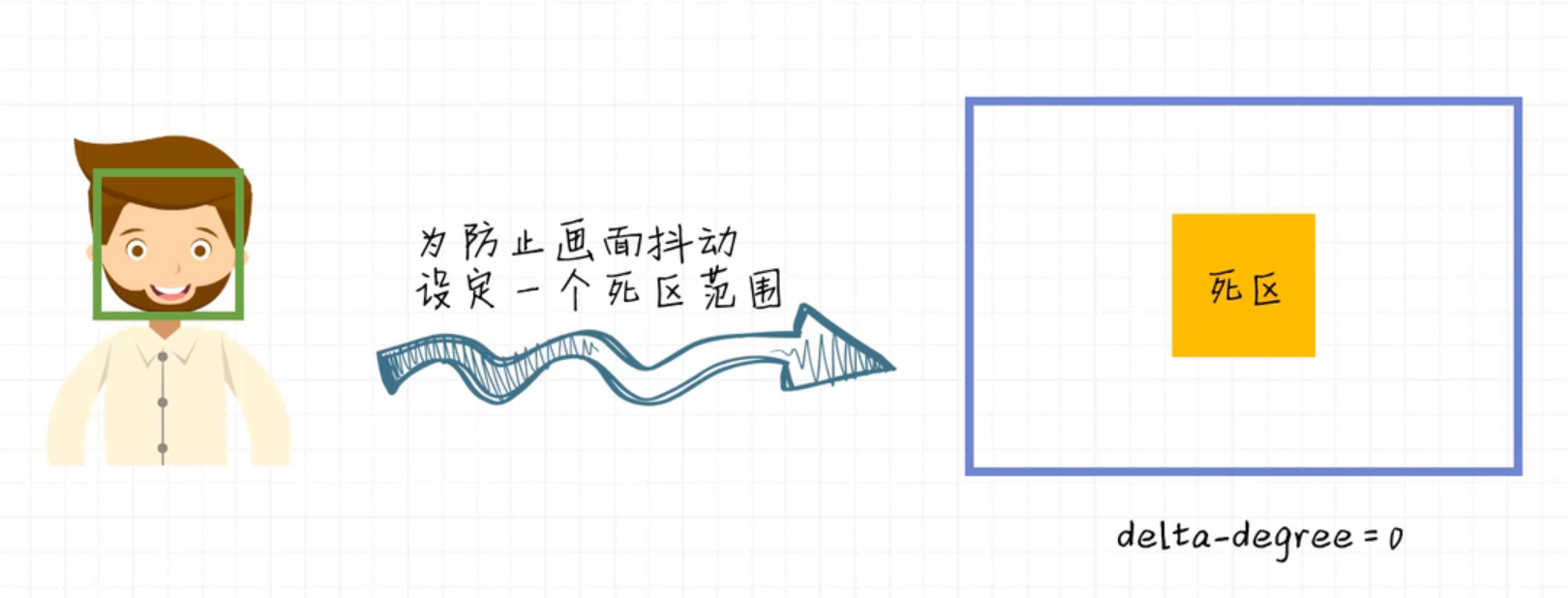
考虑到不同分辨率的因素，将偏移量整体放缩到 [ -1, 1 ]

1. 比例系数



定义一个比例系数（KP）来描述偏移量和舵机角度增量。

1. 设定死区范围



为防止画面抖动，为偏移量设置一个阈值，低于阈值不转动舵机

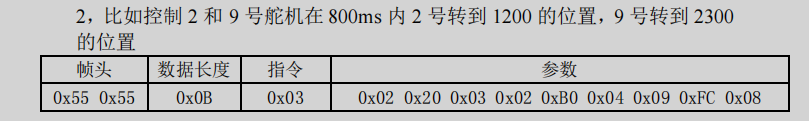
3. 采集端与舵机控制板通信

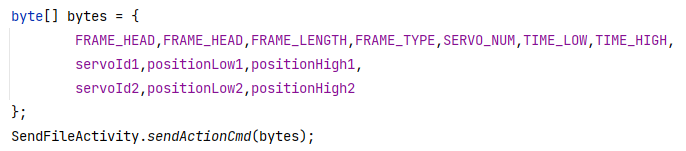
a.将得到的每个舵机角度分别转化为高位和低位数字

通过右移8位得到高位数字，与11111111相与得到低位数字。



b.依据控制板的串口通信协议编写指令





c.通过蓝牙将指令传输至控制板

## 限制条件

手机设备和舵机蓝牙已连接