

# 用户交互技术课程项目文档

——基于 Mediapipe 手势控制画图软件



小组成员：

1950059 成以恒

1951121 宁之恒

1953601 李文斐

1951606 林浩川

目录

- 1 项目背景与开发动机 ..... 3
  - 1.1 项目背景 ..... 3
  - 1.2 项目动机 ..... 3
- 2 相关工作 ..... 4
  - 2.1 技术调查 ..... 4
  - 2.2 市场调研 ..... 4
- 3 项目目标 ..... 5
  - 3.1 项目总体目标 ..... 5
- 4 概念设计 ..... 6
  - 4.1 需求分析 ..... 6
    - 1. 开始绘画 ..... 6
    - 2. 保存与再编辑 ..... 6
    - 3. 基本操作 ..... 6
    - 4. 拓展操作 ..... 6
  - 4.2 低保真度界面原型 ..... 6
  - 4.3 隐喻设计与评估 ..... 10
    - 4.3.1 设计阶段 ..... 10
    - 4.3.2 评估阶段 ..... 10
  - 4.4 交互类型 ..... 11
- 5 物理设计 ..... 11
  - 5.1 功能设计 ..... 11
  - 5.2 用户交互设计 ..... 11
  - 5.3 界面设计 ..... 11
- 6 项目实现 ..... 12
- 7 项目评估 ..... 17
- 8.团队分工与贡献 ..... 17

# 1 项目背景与开发动机

## 1.1 项目背景

随着计算机在社会中的普及，促进人机交互（HCI, Human-Computer Interaction）技术的发展将对计算机的使用产生积极影响。因此，人们越来越重视研究用于跨域人机障碍的新技术。研究的最终目标是将人机交互变得如同人与人交互一样自然。长期以来，手势一直被认为是一种可以提供更自然、更有创意和更直观的与我们的计算机进行通信的交互技术。为此，在人机交互中加入手势是一个重要的研究领域。

谈起手势识别技术，由简单粗略的到复杂精细的，大致可以分为三个等级：二维手型识别、二维手势识别、三维手势识别。

### 1. 二维手型识别

可称为静态二维手势识别，识别的是手势中最简单的一类。这种技术在获取二维信息输入之后，可以识别几个静态的手势。其缺点也很明显：拓展性差，控制感很弱，用户只能实现最基础的人机交互功能。

### 2. 二维手势识别

这一阶段仍然不含深度信息，停留在二维的层面上。这种技术不仅可以识别手型，还可以识别一些简单的二维手势动作，拥有了动态的特征。但对于复杂的人机交互，还是有一定的困难

### 3. 三维手势识别

基于有深度的信息，三维手势可以识别各种手型、手势和动作。从硬件层面接收数据传递到算法层面，其中常见的应用框架有 Microsoft Kinect，Google Mediapipe。

## 1.2 项目动机

随着如今远程办公的兴起以及 PC 使用的普及，工作者对于在 PC 端的绘画也有了更多的需求。由于画图具有更为直观且具体的表现，因此有时团队会在开会讨论时通过绘制草图讨论需求；老师也会需要绘图进行板书，便于给学生进行讲解。

基于如下几点我们决定开展本项目

- PC 端专业的画图软件上手门槛较高，学习成本高。如：photoshop 等软件，不适用于非本专业的工作者日常使用。
- 由于市面上大部分 PC 不支持触屏，所以使用一些较为简单的画图软件，如 windows 自带的画图时，只能使用鼠标操作画笔进行绘画，用户体验不佳，无法画出较为顺滑的线条。
- 若为了在 PC 端实现较为良好的绘画体验，需要在 PC 端外接绘画板等硬件进行辅助，不便捷且需要额外开销。
- 由此联系手势识别与控制技术，我们考虑到可以利用 PC 自带的前置摄像头识别用户的手势动作。将用户的手指作为画笔与 PC 进行交互，省去鼠标的交互，实现利用手势控制进行绘画。

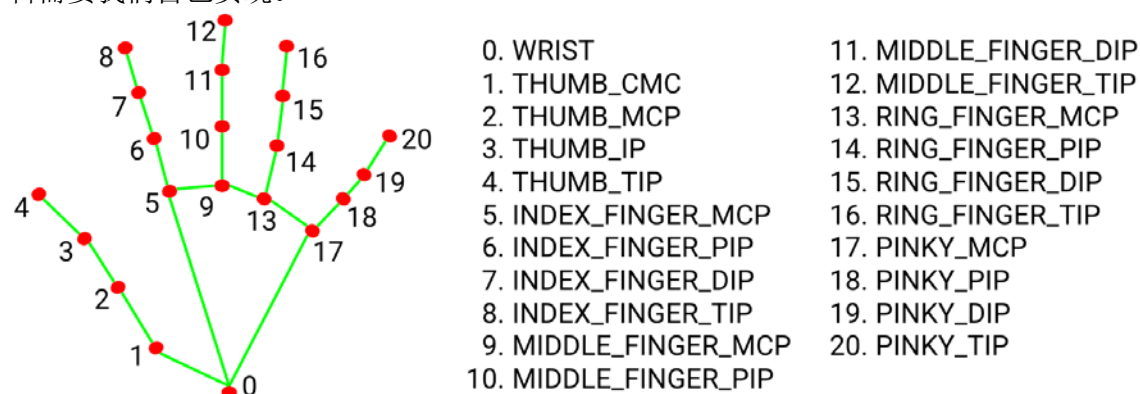
## 2 相关工作

### 2.1 技术调查

#### 手势识别与定义

需要学习 Google Mediapipe 的相关知识，实现对其内容的调用，从而达到对人体手势的准确追踪。除此之外，通过搭建框架实现对不同手势的含义进行定义，以便于后续执行过程中的调用与实现。

基于 mediapipe，我们可以得到任何一只手的 21 个关节的坐标。但需要注意的是，对于手势定义，左右手的判断及协作，手势数字识别，手指相对位置判定等相关功能逻辑需要我们自己实现。



#### GUI 的实现

研究在对三维手势识别的机制下如何做到对 GUI 实现，从而使得用户可以通过手部动作操纵屏幕上的图标或菜单选项，以选择命令、调用文件、启动程序或执行其它一些日常任务。在 GUI 的设计过程中注重如何提高用户交互体验，使得交互界面对用户更加友好。

#### 静态图像合并

通过对普通画板记录绘画操作的机制进行研究，在对转换实现手势作画时，将两幅静态图像进行合并，将用户每次创建的新绘画记录添加至原有画布中，以达到用户通过手势作画来进行二次创作的目的。

#### 素材库建立

实现在画板中插入素材库的功能，目的在于让用户可以寻找一些常用的绘图元素并插入画板，从而达到在绘图的过程中实现对素材库的调用。这一点在后续的绘画图形中有所体现。

### 2.2 市场调研

#### 绘图软件调研

调查市场上流行的绘图软件，研究其 GUI 方面的可取之处，并在设计手势识别绘板的 GUI 加以借鉴。此外研究绘板的素材库功能，收集市场上常用的素材库。

#### 应用场景调研

对手势识别绘画/书写功能进行市场调研，尝试将该技术运用到更多的应用场景之中，实现技术的多功能化。

## 3 项目目标

### 3.1 项目总体目标

通过对项目背景动机的确认与对相关技术的调研，我们最终确定了总体项目目标：**实现一个通过手势控制的画图应用软件，并拥有下述特性：**

#### 新颖、简单的作图体验

- 鼠标作图不易掌控；
- 通过手势控制绘图，让用户可以摆脱鼠标的束缚；
- 通过手指在空中描绘，作出自己想要的图像；
- 使电脑上作图、写字更容易控制。

#### 丰富的功能

- 丰富的基础功能，让用户能顺利流畅使用；
- 通过不同手势切换功能，让画图过程更流畅、简便。

#### 多元的创作素材库

- 多元的创作素材库，通过拖拽即可使用；
- 降低创作门槛，让创作更加简单、方便。

#### 多种应用场景

- 可应用于网课板书、项目汇报等场景；
- 使写字更简单、易控制；
- 增强板书的真实感。

细化目标如下所示：

1. 面对学生、教师等非专业绘画办公人士，设计一款基于手势控制的画图软件，满足此类用户在**办公场景**下对于画图软件**低成本、高效率、易操作**的需求。
2. 结合手势操作与人体姿态评估技术，从用户的习惯出发设计一款**易操作**的画图软件。
3. 确定产品形态为 PC 端应用程序，能兼容主流操作系统 MAC 以及 WINDOWS，利用笔记本自带的前置摄像头进行手势控制与姿态评估，使得开发与使用都为**低成本**。
4. 让目标用户体验到手势控制画图软件的便捷性，提高信心和对该系统的信任，今后可以主动使用该系统。

根据上述目标，接着我们确立了**可用性目标**，如下所示：

#### 1. 用户可操作性

95%的新用户可以自主操作并实现手势控制的画图。

#### 2. 高效性

90%的用户可以在 15 秒内熟悉手势控制的基本操作，并实现对于画板内容的保存。

#### 3. 系统可视化

用户在选择对应功能模块时，软件右上角会显示此操作的含义与提示，失败率低于 1%

#### 4. 准确性

用户选择对应模块以及层级时，误触其他模块的概率低于 5%

系统能快速识别并响应用户的手势以及分辨左右手，错误率低于 1%

#### 5. 一致性

97%的用户能理解软件中的各个图标的含义，不会对不同的图标理解为同种意义，也不会对一种图标产生二义性。

## 4 概念设计

### 4.1 需求分析

在需求分析阶段，根据用户的基本需求对可能的功能点进行合理分析。在该画板中，功能主要分为四个模块，即开始绘画、保存与再编辑、基本操作、拓展操作。具体内容如下：

#### 1. 开始绘画

当用户开始想要作画时需要新建一个绘图文件。

#### 2. 保存与再编辑

当用户想暂停绘画并想要保存绘画记录时需要保存绘图文件。

当用户想要在已有文件上进行再次创作时需要打开已有的绘图文件。

#### 3. 基本操作

用户想要作画需要拥有作画的画笔。

用户可以对绘图时的画笔粗细与颜色进行挑选。

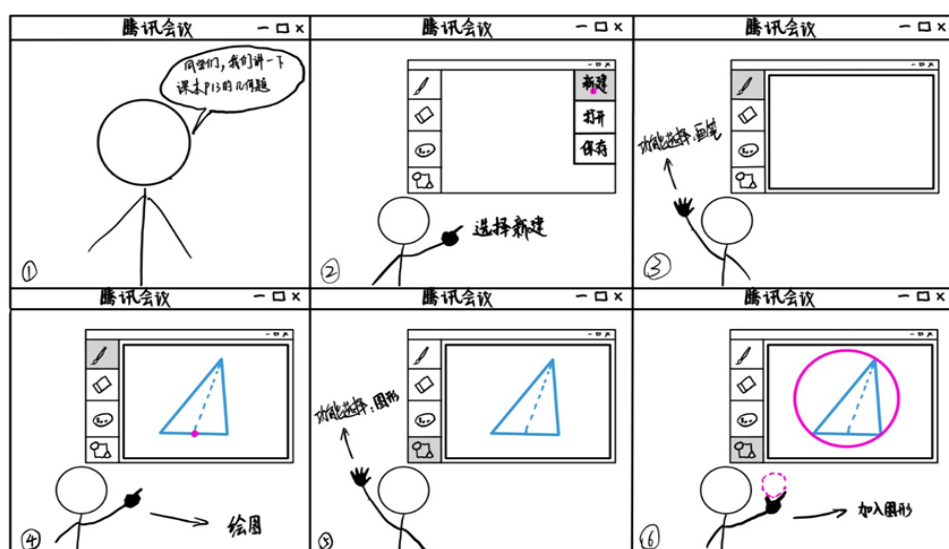
用户想要消除已有的绘画记录时，需要能做到清除。

#### 4. 拓展操作

用户可以在画布中添加系统自带图形元素。

### 4.2 低保真度界面原型

故事版模型：



卡片用例模型：  
初始界面：

画笔

橡皮

形状

颜色

新建

打开

保存

当前选择：

画笔：

画笔

橡皮

形状

颜色

新建

打开

保存

<

粗细

细

中

粗

当前选择：画笔细。

橡皮:

画  
笔

橡  
皮

形  
状

颜  
色

新  
建

打  
开

保  
存

<

粗细

细

中

粗

当前选择: 橡皮中

形状:

画  
笔

橡  
皮

形  
状

颜  
色

新  
建

打  
开

保  
存

<

形状

○

□

△

◇

☆

▽


↔

→

◎

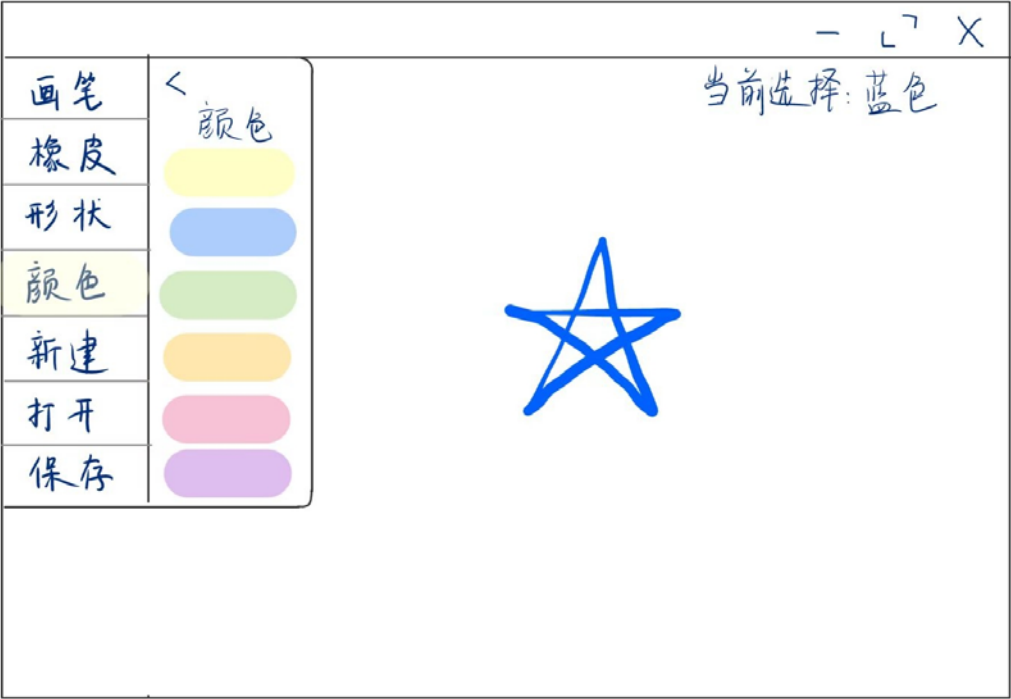
⊙

当前选择: 圆形

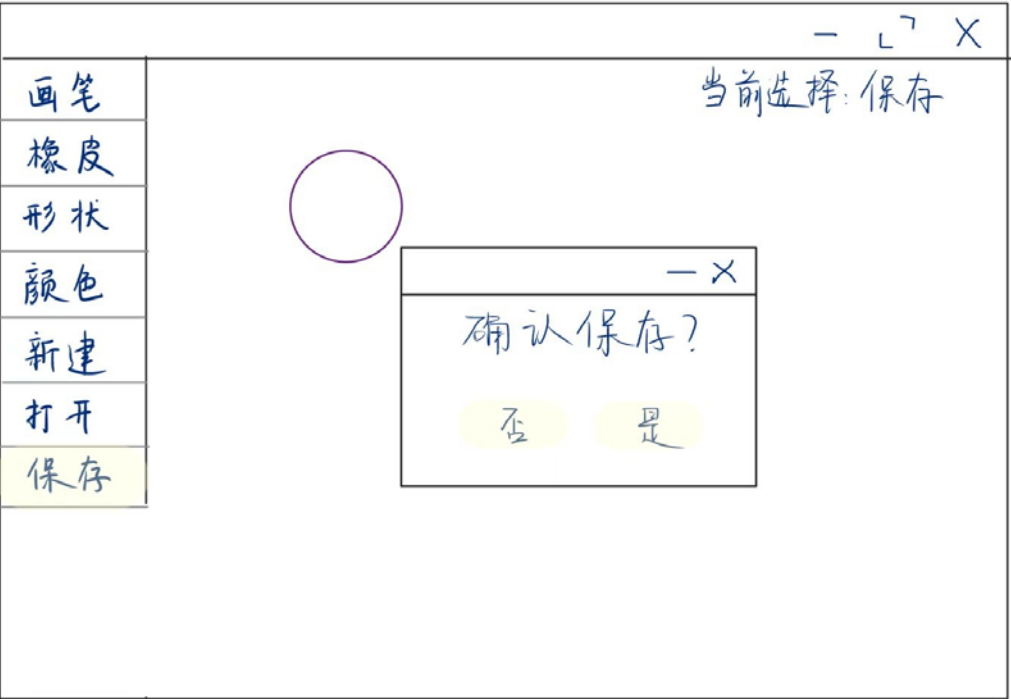




颜色：



保存：



## 4.3 隐喻设计与评估

### 4.3.1 设计阶段

根据用户需求分析，我们对各个可能功能点的 icon 进行了设计，具体解释与图标如下：

#### 保存

是一个磁盘的图形，其喻意是指将创作的内容记录进磁盘中，有保存的意思。

#### 新建

用一个文件上带一个加号代表新建，即代表了增加了一个绘图文件。

#### 打开

用一个文件夹加箭头，代表打开文件的操作，主要是对打开文件夹的动作的描述。

#### 画笔

用一支画笔的形状代表绘图的操作，是对拿画笔绘画动作的描述。

#### 橡皮

用一块橡皮在擦东西的图形代表消除绘画痕迹，是对拿橡皮擦除痕迹动作的描述。

#### 颜色

用一块调色盘来代表选择颜色，因为画家想要调出想要的颜色需要在调色盘上事先调好，故有选择颜色的意思。

#### 形状

用圆正方形和三角形的图形代表图形库，因为是由多个图案组成，故代表了图形库。



### 4.3.2 评估阶段

对于我们的隐喻，认为能够较好地满足用户的需求，基本能够实现用户对画板功能的预期。其次图标整体以简约的风格为主，彰显其功能含义，更加易于人的理解。用了很多生活中的图形物品，使得设计更加生动形象，同样也是有助于用户的理解和使用。除此之外，我们的设计也具有较好的拓展性。

## 4.4 交互类型

本画板想要实现的是不同于普通画板的绘画方式，采用的是手势识别的实现方式，对于功能的选择和切换以及绘画的方式均由手势动作实现。

其基本操作是伸出手指之后，摄像头捕捉到手指，则对左右手以及手势进行识别，对其含义进行解析处理。对于基本的绘画操作，则直接可以通过食指在画布上进行绘图。在概念设计阶段，对功能切换操作的初步构想的是用两只手来实现不同的功能，而对于具体的某一只手，则根据该只手的手势对其功能下的细节进行挑选。

# 5 物理设计

## 5.1 功能设计

**功能：**

- 画笔：选择画笔粗细；
- 橡皮：选择橡皮粗细；
- 形状：选择某个形状进行绘制；
- 颜色：选择绘制颜色；
- 新建：新建画布；
- 保存：保存绘制的内容；
- 打开：打开一张图片，在该图片基础上进行绘制。

**功能关联性：**

用户选择画笔、形状、颜色进行绘图；对于绘图有误的地方，可以选择橡皮擦除；绘制完成后，可以保存图片；可以新建画布，重新开始绘制。

## 5.2 用户交互设计

**用户交互方式：**

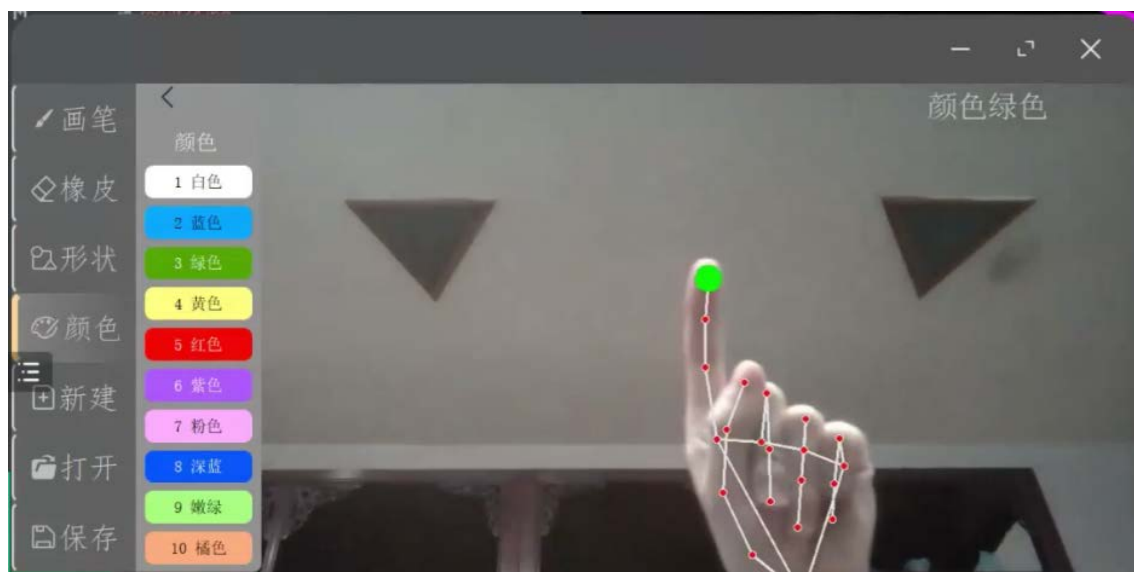
用户通过手势交互选择功能并进行绘图。用户通过左手位置，选择基本功能，然后根据手势选择详细的功能，如画笔粗细、颜色等，再通过右手食指的位置进行绘图。

**用户交互场景：**

对于同学开会、老师讲课等办公场景下，用户需要绘图，但又不适应鼠标的操作，可以通过手势操作，使用本软件，进行简单的绘图操作。

## 5.3 界面设计

- 软件 UI 为简约风，以功能为导向，突出功能，方便用户使用。
- 左边为一级菜单，右边为画板，方便用户左手选择功能、右手进行绘画。
- 一级菜单为列表式导航，每个按钮较高，方便用户定位，增加了选择的容错性。
- 二级菜单为抽屉式，根据用户选择的不同功能而展开不同的二级菜单，让用户通过手势选择，方便用户操作，减少了选择错误的可能性。
- 右上角有选择提示，用户可以直观的看到自己的选择。

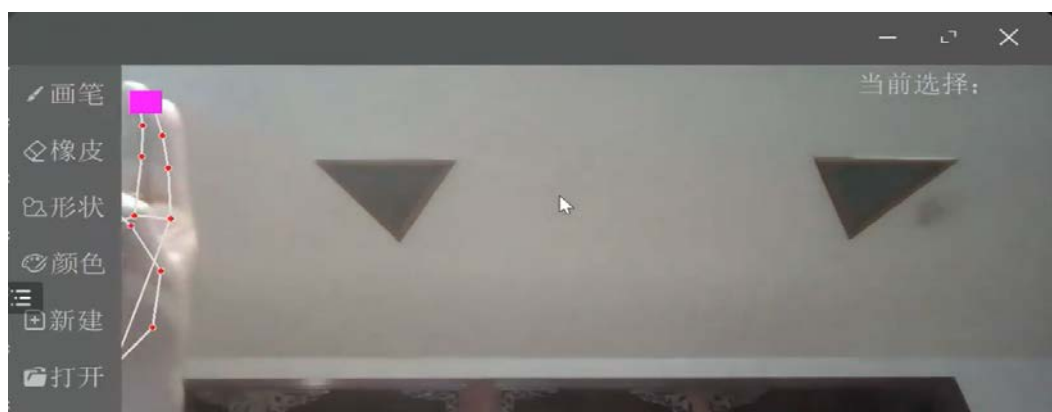
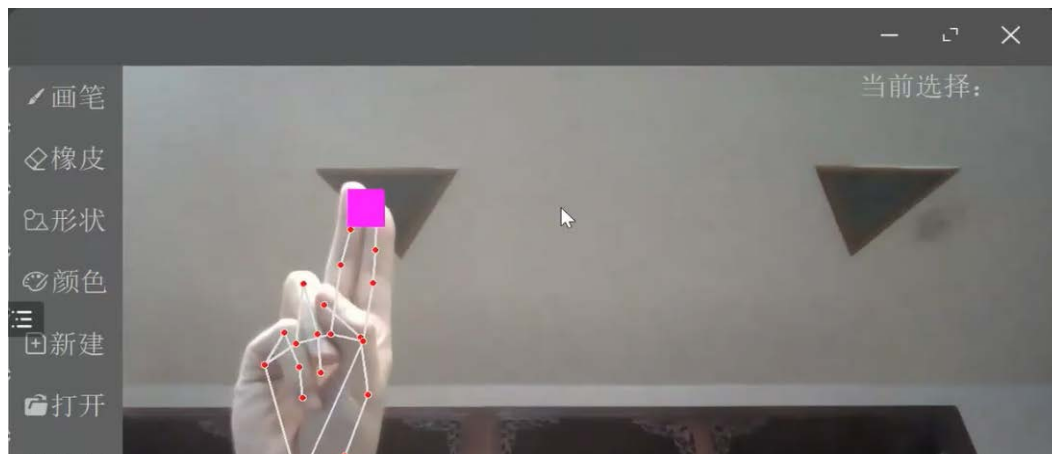


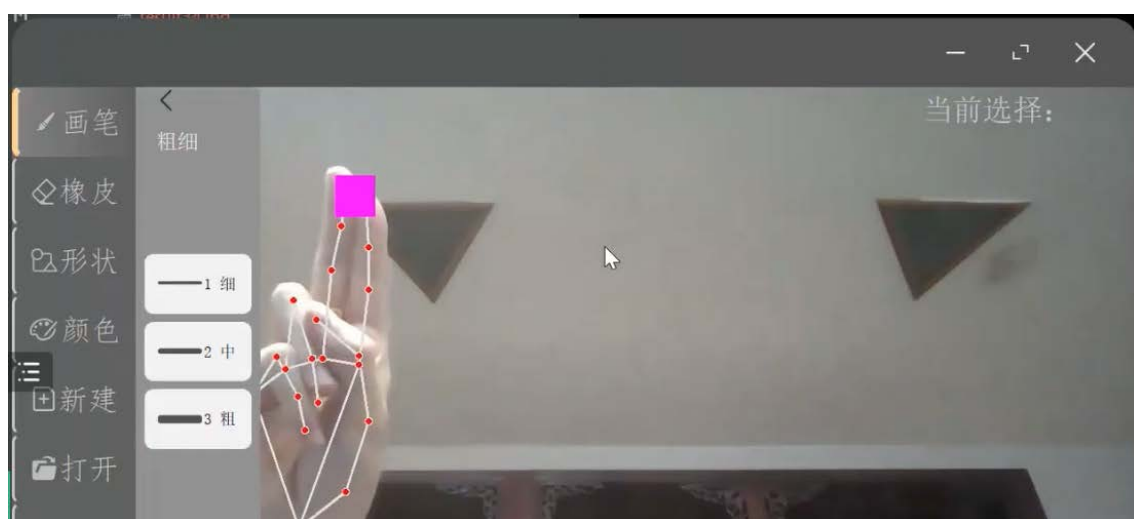
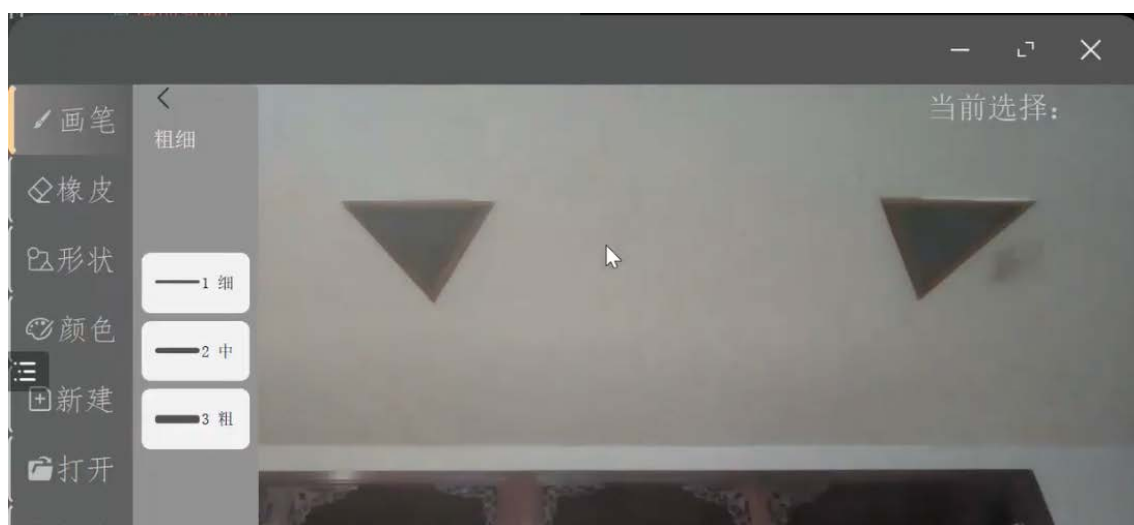
## 6 项目实施与结果展示

根据之前对于项目的概念设计与物理设计，我们最终设计出一套较为简洁且易操作的手势控制逻辑：左手负责选择工作，右手负责绘制工作

下述我将着重阐述我们对于手势交互逻辑的定义，其余功能点以及操作演示可以查看我们项目的演示视频

### 1. 一级菜单栏选择



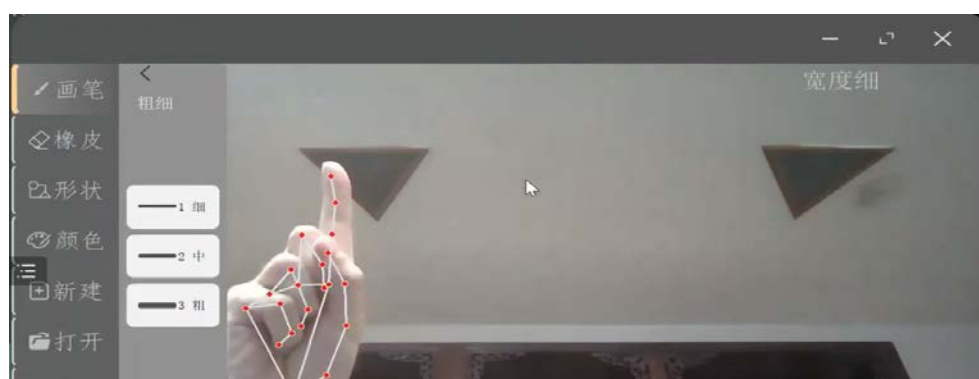


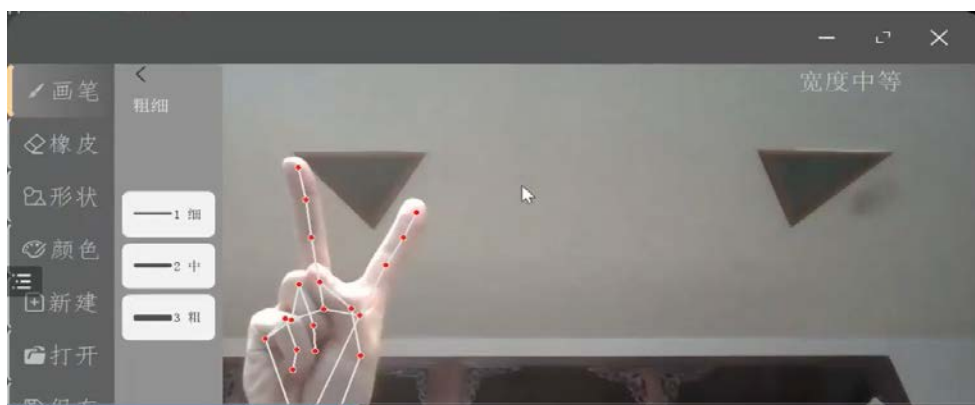
考虑到用户的可操作性，我们项目使用**隔空触控式交互形式**进行第一层级的选择。即当用户左手食指与中指并起并抬起时，判定进入第一层级选择状态。用户的食指处将会生成一个提示色块，此色块会一直跟随用户左手食指运动。用户可以将手指移动到目标位置选择，如画笔、橡皮、形状等。

当手指被判定到达目标区域时，会展开二级菜单栏。如图所示。

综上，一级菜单栏选择成功。

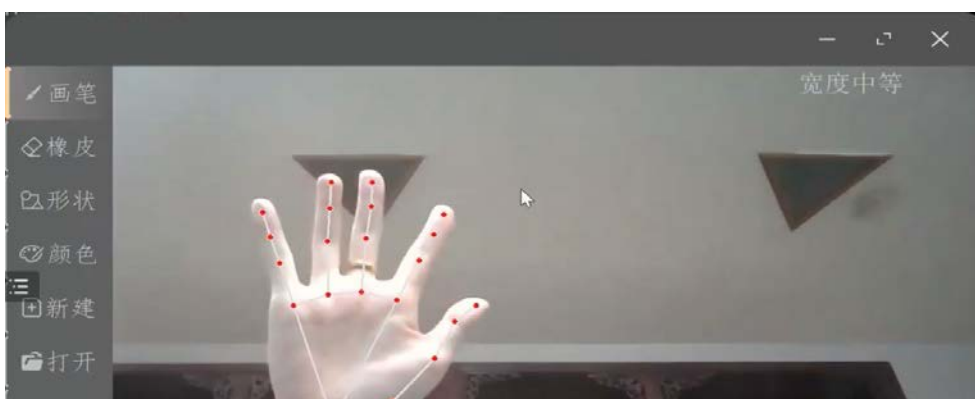
## 2. 二级菜单栏选择





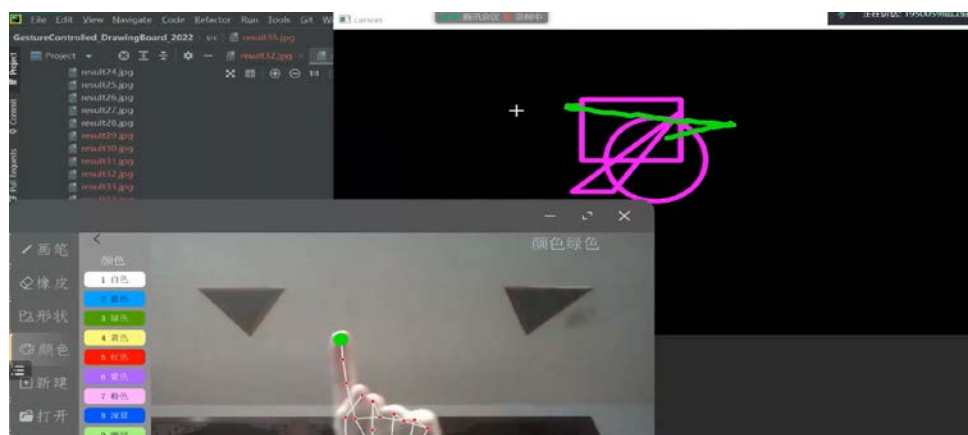
考虑到便捷性与准确性，我们项目使用**手势语义控制**进行第二层级的选择。即我们根据手指的姿态进行识别，并定义其对应的手势语义。如：当用户用手指比划出阿拉伯数字 1 时，项目会判定用户选择此二级菜单的第一个子属性；而用户比划出阿拉伯数字 2 时，项目会判定用户选择了第二个子属性。以此类推，如图所示。对于**手势语义控制**，我们认为其能大大提升用户选择的效率以及正确性。此交互逻辑具有**准确、便捷、高效**等特性。

### 3. 二级菜单栏退出



对于二级菜单栏的退出，我们将伸出五指的状态定义为二级菜单栏退出控制。如图所示，其简单易懂，准确性高，适合应用于菜单退出等此类频繁且高需求量的操作。

### 4. 绘图



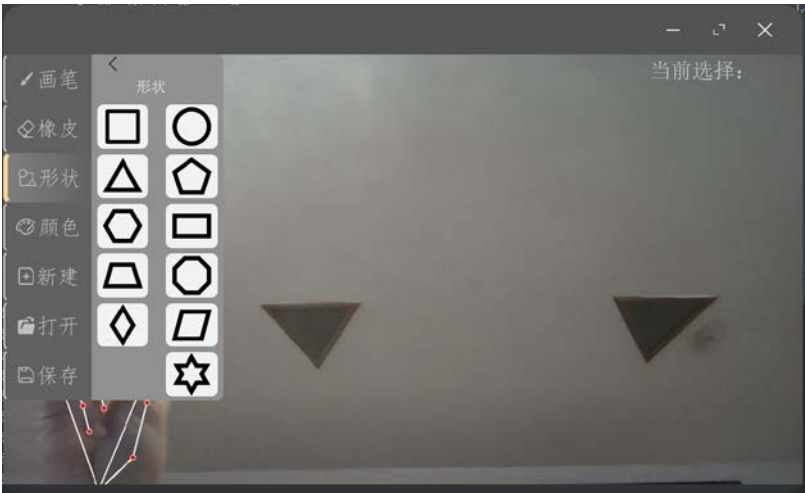
伸出右手的食指作为画笔，系统会追踪食指的运动轨迹并图像显示在绘图板上。



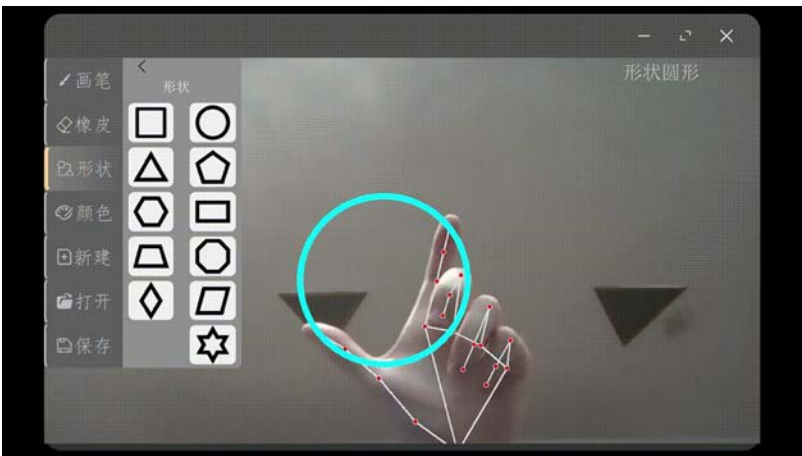
以下为项目运行截图：



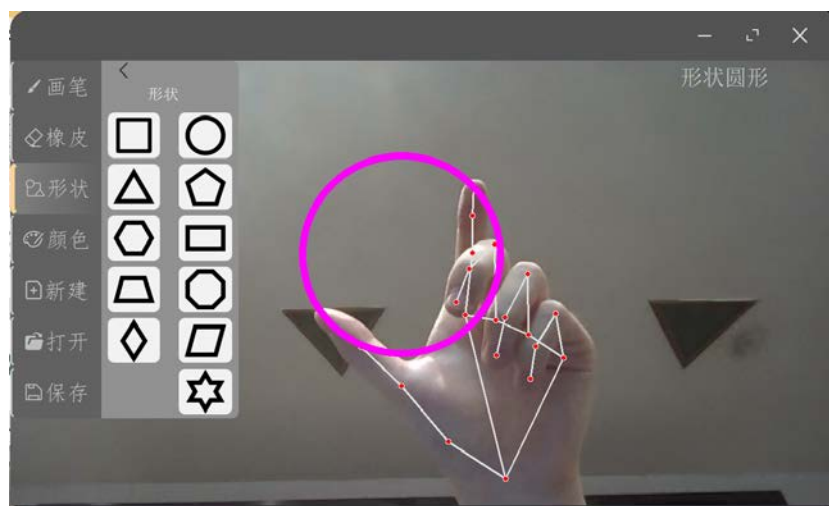
左手的识别



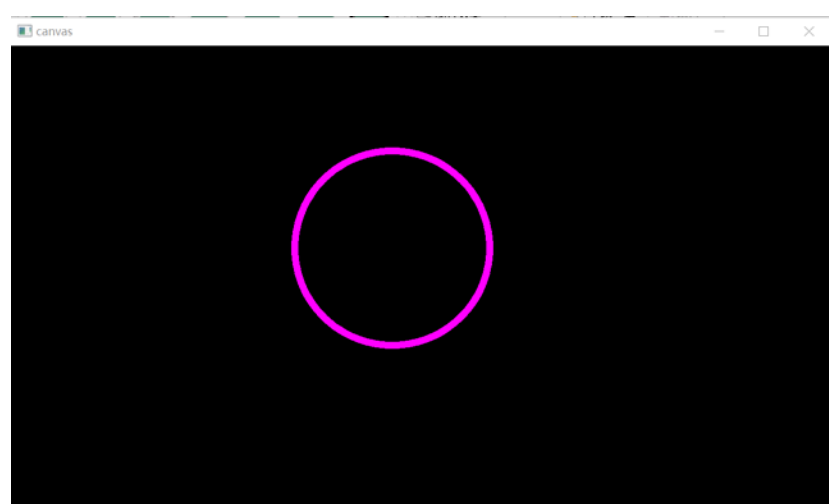
二级菜单栏的展开



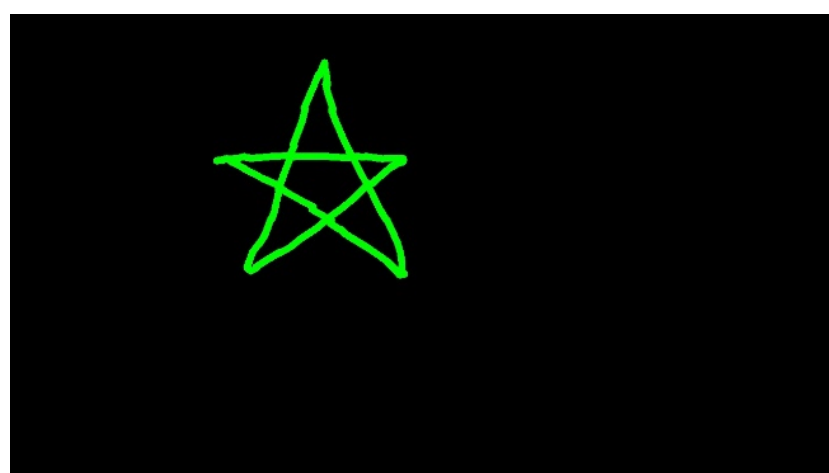
圆形绘制过程



圆形绘制过程



圆形绘制结果



手绘五角星结果



## 7 项目评估

我们采用了可用性测试进行评估，根据之前列出的一些可用性目标，我们邀请了同学及家长作为用户使用本产品，测试过程中通过与用户的沟通，我们按照架构、布局、内容以及交互四个大点进行整合，如下所示：

### 架构与导航

- 控制面板采用了左侧导航栏与全屏摄像头手势识别，用户易于理解与熟悉，页面结构清晰
- 能通过手势快速退出选择栏，回到首页
- 对于手指使用色块跟踪处理，能识别其在页面的位置

### 布局与设计

- 导航栏有两层，一层主控制栏，一层副控制栏，界面元素易懂
- 文字、图标形状美观且大小合适，容易识别，不会出现二义性
- 主要操作区域面积大，视线流畅

### 内容与可读性

- 提示语言易懂、简洁、礼貌
- 内容表达意义一致，无二义性
- 重要的导航栏与当前状态显示栏信息都位于显著位置

### 行为与交互

- 用色块跟踪用户的手指，引导用户进行选择
- 最基本的选择、退出、绘制操作简单易懂
- 有对当前用户画笔的状态提示
- 有对用户进行保存操作时的行为反馈与弹窗
- 已将操作步数化为最简，且左右手各司其职，其各自作用分离，易于操作与理解

综上，可以看到我们项目总体可交互性强，用户体验良好。

## 8.团队分工与贡献

成以恒：查阅资料，底层框架搭建，管理与运维 25%

林浩川：需求分析，UI 框架搭建，UI 样式设计 25%

李文斐：需求分析，UI 原型设计，UI 接口设计 25%

宁之恒：需求分析，底层框架搭建，手势捕捉与定义 25%