

به نام خدا



## پروپوزال

آز سخت افزار-گروه ۳- دکتر اجلالی

دانشکده مهندسی کامپیوتر

دانشگاه صنعتی شریف

نیم سال دوم ۰۱-۰۰

نویسندگان:

علی حاتمی تاجیک-۹۸۱۰۱۳۸۵

امیر محمد عیسی زاده-۹۸۱۰۶۸۰۷

محمد حسین فیصربه-۹۷۱۰۶۲۳۸



## فهرست مطالب

۲	۱ هدف
۲	۲ بررسی طراحی
۲	۱.۲ طراحی با پردازش روی ماشین مقصد
۲	۲.۲ طراحی با پردازش مجزا روی محصول
۳	۳ بررسی جزئیات
۳	۱.۳ بستر پردازشی
۳	۲.۳ تصویر
۳	۳.۳ پردازش تصویر
۳	۴.۳ اتصال به سیستم
۴	۵.۳ Driver
۴	۴ تخمین هزینه
۴	۵ زمانبندی



## ۱ هدف

هدف این پروژه ساخت کنترلی است که با استفاده از حالت‌ها<sup>۱</sup> و حرکات دست بتوان یک رایانه را کنترل کرد. این عملیات به کمک پردازش تصویر گرفته شده از دست کاربر که مکان و جابجایی‌های آن به حرکت نشانگر موشواره و عملیات‌های مربوط به آن منجر می‌شود. در کنار این تعدادی حالات قابل تعیین برای کاربر برای روانتر شدن داشته رابط کاربری وجود داشته باشد. در این پروژه نیازی به مازول شدت صوت نیست و تمام حالت‌ها به وسیله پردازش تصویر قابل دستیابی هستند.

## ۲ بررسی طراحی

برای این پروژه دو طراحی<sup>۲</sup> کلی می‌توان متصور بود که در ادامه بررسی خواهند شد.<sup>۳</sup>

### ۱.۲ طراحی با پردازش روی ماشین مقصد

در این طراحی محصول نهایی تنها تصویر را از محیط جمع‌آوری کرده و برای ماشین مقصد ارسال می‌کند. بار پردازش این داده‌ها بر روی ماشین مقصد خواهد بود که به همین خاطر دستگاه تنها روی ماشین‌هایی که قدرت پردازشی بالایی دارند (مثل رایانه‌های شخصی و سرورها) قابل استفاده هستند. از طرفی ارسال این اطلاعات به رایانه باعث می‌شود تا محصول نهایی باریک و سبک باشد. همچنین می‌توان نرم‌افزار آنرا به تنهایی به فروش رساند و از سخت‌افزار آن تنها یک بخش انتخابی<sup>۴</sup> باشد و هر کاربری با استفاده از یک دوربین معمولی و پیکربندی<sup>۵</sup> مناسب آن بتواند از قابلیت‌های آن استفاده کند. مزایا:

- وزن و ابعاد محصول نهایی کوچکتر است.
- هزینه تمام شده برای محصول نهایی کمتر است.
- امکان فروش به صورت نرم‌افزار صرف را دارد.
- می‌تواند به صورت بی‌سیم مورد استفاده قرار بگیرد.

معایب:

- تنها روی ماشین‌های قدرتمند قابل استفاده است (برای مثال قابل استفاده بر روی تلوزیون‌های هوشمند و کنسول‌های بازی نیست).
- مقدار قابل توجهی از منابع کاربر را به صورت مداوم مشغول می‌کند.

### ۲.۲ طراحی با پردازش مجزا روی محصول

در این نوع طراحی باید یک هسته پردازشی قدرتمند با مقدار کافی حافظه در کنار دوربین قرار بگیرد تا پردازش به صورت بی‌درنگ<sup>۶</sup> صورت بگیرد و نتایج حاصل از پردازش تصاویر با پروتکل‌های خاص<sup>۷</sup> به ماشین مقصد ارسال شود و در آنجا با استفاده از یک درایور تفسیر شوند. در این طراحی کاربر لزوماً باید سخت‌افزار را خریداری کرده که این می‌تواند ارزش افزوده زیادی ایجاد کند. مزایا:

- ماشین مقصد نیاز به قدرت پردازشی (زیادی) ندارد.
- می‌توان آنرا به تمام ماشین‌هایی که به موشواره و صفحه‌کلید نیاز دارند متصل کرد.

Gestures<sup>۱</sup>

Design<sup>۲</sup>

آلبته طراحی‌های هوشمندانه‌تر و فناوری‌ها نیز قابل تصور است، برای مثال با استفاده از یک فضای خازنی و بدون استفاده از دوربین پردازش‌ها انجام بگیرد تا بسته به تغییر شرایط محیط، نور و ... کارایی بهتری داشته باشد اما از آنجایی که در صورت پروژه کارفرما درخواست ساخت مازول با استفاده از تکنولوژی پردازش تصویر را داشته است به طراحی‌های منحصر به همین حوزه بسنده شده است.

Optional<sup>۴</sup>

Configuration<sup>۵</sup>

Real-Time<sup>۶</sup>

Human Interface Device (HID)<sup>۷</sup>



- ارزش افزوده بیشتر برای محصول نهایی متصور است.
- انحصار نسبی محصول برای شرکت سازنده حفظ می‌شود.

معایب:

- به خاطر پردازشی که انجام می‌دهد استفاده از باتری به صرفه نیست و لزوماً باید از کابل برای اتصال استفاده شود.
- وزن و هزینه محصول نهایی بیشتر خواهد بود.

## ۳ بررسی جزئیات

### ۱.۳ بستر پردازشی

- طراحی اول: در طراحی اول نیازی به یک بستر پرقدرت برای پردازش نداریم و یک برستر پردازشی به اندازه‌ای که بتواند تصاویر را از سنسور دریافت کرده و آنها را به صورت دیجیتال به کامپیوتر مقصد ارسال کند کفایت می‌کند. برای این کار می‌توان از برد ESP32 که دارای ارتباطات Wi-Fi و بلوتوث است استفاده کرد. انرژی این برد را می‌توان هم به وسیله باتری و هم به وسیله اتصال مستقیم به رایانه مقصد تامین کرد (به خاطر محدودیت‌های باتری از کابل برای توان بخشی این ماژول استفاده خواهد شد).
- طراحی دوم: در این طراحی نیاز به قدرت پردازشی زیادی داریم و بردهای خانواده آردوینو<sup>۸</sup> پاسخگوی این قدرت پردازشی نخواهند بود و گزینه مناسب استفاده از خانواده رزبری پای<sup>۹</sup> است.

البته تمام اعضای این خانواده مناسب کار نیستند. همانطور که بالاتر اشاره شد تامین انرژی این طراحی نمی‌تواند به وسیله باتری انجام شود و حتماً باید به یک منبع تغذیه خارجی متصل باشد و ارتباطات باید به وسیله سیم انجام بگیرد. مشکلی که در اینجا به وجود خواهد آمد این است که در برخی خانواده‌ها پورت‌های USB آن از قابلیت OTG<sup>۱۰</sup> پشتیبانی نمی‌کنند (به خاطر هایی که وجود دارد و همینطور استفاده از آن درگاه‌ها به عنوان درگاه اترنت) و تنها اعضای خاصی این قابلیت را دارند که می‌توان از RPi Zero نام برد که در کنار قدرت پردازشی مناسبی که دارد اندازه کوچکی نیز دارد و مصرف توان آن هم خیلی زیاد نیست.<sup>۱۱</sup>

### ۲.۳ تصویر

برای دریافت تصویر در طراحی اول از ماژول دوربین که همراه با برد ESP32 ارائه می‌شود می‌توان بهره برد. می‌توان فیلتر مادون قرمز آنرا نیز جدا کرد تا دوربین دید در شب داشته باشیم و با یک سورتس مادون قرمز کاربری دستگاه را در هنگام نبود نور مرئی نیز حفظ کنیم. در طراحی دوم باید از ماژول‌های دوربین طراحی شده برای خانواده رزبری پای استفاده کرد. برای این کار می‌توان از حسگر ۵ مگاپیکسلی OV5647 NoIR استفاده کرد که به همراه چراغ‌های مادون قرمز هم ارائه می‌شود.

### ۳.۳ پردازش تصویر

برای پردازش تصویر دریافتی از قسمت قبل از یک برنامه تحت زبان برنامه نویسی پایتون و با استفاده از کتابخانه‌های OpenCV2 و MediaPipe استفاده می‌شود که این کتابخانه‌ها با استفاده از شبکه‌های عصبی که از قبل روی تعداد زیادی تصاویر تمرین داده شده است حالت‌های دست تشخیص داده می‌شوند. این حالت‌ها در گزارش هفته اول بررسی خواند شد.

### ۴.۳ اتصال به سیستم

همانطور که در طراحی‌ها هم گفته شد در طراحی اول اتصال به سیستم از طریق وایفای انجام می‌گیرد و تصاویر به رایانه منتقل می‌شوند. در طراحی دوم این اتصال از طریق یک کابل نوع C استفاده می‌شود تا اطلاعات و توان از طریق آن منتقل شود.

Arduino<sup>۸</sup>  
RPi-Raspberry Pi<sup>۹</sup>  
On The Go<sup>۱۰</sup>

<sup>۱۱</sup> به خاطر اینکه بردی که دانشگاه قادر است در اختیار دانشجویان قرار دهد یک از قابلیت OTG پشتیبانی نمی‌کند و خرید یک برد رزبری پای جداگانه هزینه‌های ساخت را بالا می‌برد ترجیح بر این است که از طراحی اول استفاده شود. البته یک راه دیگر وجود دارد که یک سرور روی شبکه محلی ایجاد شود که هر رزبری پای بتواند دستورات دیکت شده را به آن بفرستد ولی این کار به خاطر قابل اتکا نبودن می‌تواند منجر به تجربه نامطلوب کاربران بشود.



### ۵.۳ Driver

با استفاده از کتابخانه pyautogui می توان تغییرات موسواره و کیبورد را به رایانه منتقل کرد. در طراحی اول چون پردازش روی همان رایانه انجام می شود این ماژول مستقیماً به پردازش تصویر متصل می شود و فرامین را اجرا میکند.

در طراحی دوم اما، از آن جهت که پردازش روی دستگاه انجام می شود سیگنال های پردازش شده از طریق USB به رایانه منتقل می شود و با استفاده از پروتکل های استاندارد جهانی (مثل HID) این سیگنال ها مانند سیگنال های trackpad به رایانه ها منتقل خواهند شد.

### ۴ تخمین هزینه

- طراحی اول:

ردیف	پارت نامبر	هزینه (ریال)
۱	ESP32-Camera	۲,۰۰۰,۰۰۰
۲	3D Printed Case	۲,۵۰۰,۰۰۰
۳	Micro USB Cable	۱۵۰,۰۰۰

که با در نظر گرفتن حاشیه امن هزینه تمام شد ۵,۰۰۰,۰۰۰ ریال خواهد بود.

- طراحی دوم:

ردیف	پارت نامبر	هزینه (ریال)
۱	Raspberry Pi Zero W	۱۶,۴۵۰,۰۰۰
۲	Raspberry Pi Camera Module (NoIR) with 5MP OV5647 sensor	۵,۵۰۰,۰۰۰
۳	3D Printed Case	۳,۵۰۰,۰۰۰
۴	Micro USB to Type-C Cable	۵۰۰,۰۰۰

که با در نظر گرفتن حاشیه امن قیمت نهایی ۲۸,۰۰۰,۰۰۰ ریال خواهد بود.

### ۵ زمانبندی

زمان	برنامه
هفته اول	طراحی حالات و مقدمات پردازش تصویر
هفته دوم	پیاده سازی پردازش تصویر
هفته سوم	تهیه قطعات و پیکربندی کردن آنها
هفته چهارم	اتصال به سیستم و تست محصول نهایی