



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
دانشکده مهندسی کامپیوتر

گزارش پروژه کارشناسی

طراحی و پیاده سازی سامانه ردیاب تصویری سه بعدی چشم

نگارش
یگانه کردی

استاد راهنما
جناب آقای دکتر رحمتی

مهر ۱۴۰۰

صفحه فرم ارزیابی و تصویب پایان نامه - فرم تأیید اعضاء کمیته دفاع

در این صفحه (هر سه مقطع تحصیلی) باید فرم ارزیابی یا تأیید و تصویب پایان نامه/رساله موسوم به فرم کمیته دفاع برای ارشد و دکترا و فرم تصویب برای کارشناسی، موجود در پرونده آموزشی را قرار دهند.

نکات مهم:

- ✓ نگارش پایان نامه/رساله باید به **زبان فارسی** و بر اساس آخرین نسخه دستورالعمل و راهنمای تدوین پایان نامه های دانشگاه صنعتی امیرکبیر باشد. (دستورالعمل و راهنمای حاضر)
- ✓ تحویل پایان نامه به زبان انگلیسی، برای دانشجویان بین الملل با شرایط دستورالعمل حاضر بلامانع است و داشتن صفحه عنوان فارسی به همراه چکیده مبسوط فارسی، ۳۰ صفحه برای پایان نامه ارشد و ۵۰ صفحه برای رساله دکترا در ابتدای آن الزامی است.
- ✓ دریافت پایان نامه کارشناسی و کارشناسی ارشد، **بصورت نسخه الکترونیکی** مطابق راهنما و دستورالعمل حاضر می باشد.
- ✓ رنگ جلد رساله دکترا باید "سفید" رنگ و اطلاعات مندرج "زرکوب" باشد. چاپ و صحافی رساله بصورت **پشت و رو (دورو)** الزامی است.
- ✓ در صورتی که یک عنوان پایان نامه دارای **دو نویسنده** است، فقط یکبار فایل و فرم اطلاعات آن با ذکر هر دو نویسنده بارگذاری و تکمیل گردد.
- ✓ با توجه به اینکه در ورود ۲۰۱۶ یا بالاتر، احتمال تغییر ترتیب ذکر زیر فصل ها وجود دارد لطفا در انتها به شماره دهی زیر فصل ها توجه نمایید که بصورت صحیح باشد.
از راست به چپ: شماره فصل - زیرفصل ۱ - زیرفصل ۲ - زیرفصل ۳ و ...



به نام خدا

تاریخ:

تعهدنامه اصالت اثر

اینجانب یگانه کردی متعهد می‌شوم که مطالب مندرج در این پایان نامه حاصل کار پژوهشی اینجانب تحت نظارت و راهنمایی اساتید دانشگاه صنعتی امیرکبیر بوده و به دستاوردهای دیگران که در این پژوهش از آنها استفاده شده است مطابق مقررات و روال متعارف ارجاع و در فهرست منابع و مآخذ ذکر گردیده است. این پایان نامه قبلاً برای احراز هیچ مدرک هم‌سطح یا بالاتر ارائه نگردیده است.

در صورت اثبات تخلف در هر زمان، مدرک تحصیلی صادر شده توسط دانشگاه از درجه اعتبار ساقط بوده و دانشگاه حق پیگیری قانونی خواهد داشت.

کلیه نتایج و حقوق حاصل از این پایان نامه متعلق به دانشگاه صنعتی امیرکبیر می‌باشد. هرگونه استفاده از نتایج علمی و عملی، واگذاری اطلاعات به دیگران یا چاپ و تکثیر، نسخه‌برداری، ترجمه و اقتباس از این پایان نامه بدون موافقت کتبی دانشگاه صنعتی امیرکبیر ممنوع است. نقل مطالب با ذکر مآخذ بلامانع است.

یگانه کردی

امضا

اکنون که به یاری خداوند این دوره را به پایان رسانیده‌ام، بر خود واجب می‌دانم از اساتید راهنمای بزرگواری جناب آقای دکتر رحمتی به پاس زحمات بی‌شائبه‌شان در طی انجام این تحقیق سپاسگزاری نمایم.

همچنین از اساتید گرانمایه، جناب آقای دکتر صفابخش که زحمت داوری این پایان‌نامه را بر عهده داشتند نهایت تشکر را دارم.

از سایر اساتید بزرگ که شاگردی محضرشان از بزرگترین افتخارات زندگی علمی‌ام می‌باشد، کمال تشکر را دارم.

چکیده

در این پروژه می‌خواهیم سیستمی جهت ردیابی سه‌بعدی سمت نگاه پیاده‌سازی کنیم. این سیستم با دریافت تصاویر از چشم فرد توانایی تعیین سمت نگاه فرد در هر لحظه را خواهد داشت. و می‌توان از آن برای کاربردهای مختلف پزشکی و علمی استفاده کرد.

برای پیاده‌سازی این سیستم، از یک الگوریتم ردیابی حرکات چشم سه‌بعدی استفاده کرده‌ایم که تنها طبق مختصات هندسی مردمک عمل می‌کند. این الگوریتم نیاز به واسنجی برای هر کاربر ندارد، همچنین نیاز به تنظیم کردن نورهای محیطی ندارد. این الگوریتم براساس اعمال یک مدل سه‌بعدی برای چشم بر روی مجموعه‌ای از عکس‌های مردمک چشم عمل می‌کند. پس از اعمال مدل و بهینه‌سازی آن به نحوی که بر روی عکس‌های اصلی منطبق شود، می‌توانیم زاویه ی نگاه را بیابیم.

همچنین ما یک سخت‌افزار کم‌هزینه مشابه عینک ساخته‌ایم که استفاده از آن راحت است و می‌تواند در تحقیقات بعدی برای ردیابی سمت نگاه مورد استفاده قرار گیرد.

مراحل مربوط به آزمایش دقت پیاده‌سازی به علت کمبود منابع سخت‌افزاری در این پروژه انجام نشده است. جهت سنجش دقت الگوریتم در آینده می‌توانیم از افراد بخواهیم به چند نقطه با زوایای مشخص نگاه کنند و سپس با مقایسه نتایج الگوریتم و زاویه مورد انتظار دقت الگوریتم را محاسبه کنیم.

واژه‌های کلیدی:

ردیابی سمت نگاه، ردیاب حرکات چشم، تشخیص مردمک، مدل‌سازی چشم

فصل اول مقدمه.....	۱
فصل دوم بررسی مطالعات انجام شده.....	۴
۱-۲- انواع الگوریتم‌ها.....	۵
۲-۲- الگوریتم‌های تشخیص سمت نگاه.....	۵
۳-۲- روش‌های واسنجی.....	۶
۴-۲- جمع‌بندی فصل.....	۷
فصل سوم الگوریتم ردیابی سمت نگاه.....	۸
۱-۳- فرض‌ها و محدودیت‌ها.....	۹
۲-۳- شرح الگوریتم.....	۹
۱-۲-۳- بدست آوردن موقعیت مردمک چشم.....	۱۰
۲-۲-۳- بدست آوردن مکان هندسی مردمک.....	۱۲
۳-۲-۳- بدست آوردن مدل چشم.....	۱۴
۴-۲-۳- تخمین مرکز کره چشم.....	۱۴
۵-۲-۳- تخمین شعاع کره.....	۱۶
۶-۲-۳- تخمین مردمک پایدار.....	۱۸
۳-۳- جمع‌بندی فصل.....	۱۹
فصل چهارم پیاده‌سازی الگوریتم.....	۲۰
۱-۴- پیاده‌سازی بخش سخت‌افزاری.....	۲۱
۱-۱-۴- محدودیت‌ها.....	۲۱
۱-۱-۱-۴- نور خورشید.....	۲۱
۲-۱-۱-۴- حرکات غیرعادی چشم.....	۲۲
۳-۱-۱-۴- سیستم مختصات نسبی.....	۲۲
۲-۱-۴- دوربین‌ها.....	۲۲
۳-۱-۴- ساختار سخت‌افزار.....	۲۴
۴-۱-۴- مشکلات سخت‌افزار.....	۲۵
۲-۴- پیاده‌سازی بخش نرم‌افزاری.....	۲۶
۱-۲-۴- کتابخانه‌های مورد استفاده.....	۲۶
۲-۲-۴- واسنجی.....	۲۶
۳-۲-۴- پیاده‌سازی کد.....	۲۶
۴-۲-۴- رابط کاربری.....	۲۷
۵-۲-۴- آزمایش عملکرد پروژه.....	۲۷

۲۷.....	۳-۴- نتایج بدست آمده.....
۳۱.....	۴-۴- مشکلات و محدودیت‌ها.....
۳۲.....	۴-۵- جمع‌بندی فصل.....
۳۳.....	فصل پنجم جمع‌بندی و پیشنهادات.....
۳۶.....	منابع و مراجع.....

شکل ۱-۳ معماری شبکه عصبی.....	۱۱
شکل ۲-۳ بدست آوردن مردمک.....	۱۳
شکل ۳-۳ بدست آوردن مرکز کره.....	۱۶
شکل ۴-۳ تخمین شعاع کره چشم.....	۱۸
شکل ۱-۴ دوربین مورد استفاده.....	۲۳
شکل ۲-۴ سخت افزار طراحی شده.....	۲۵
شکل ۳-۴ خروجی الگوریتم ناحیه بندی مردمک.....	۲۸
شکل ۴-۴ حرکت مردمک در تصاویر متوالی.....	۲۸
شکل ۵-۴ نتایج الگوریتم.....	۲۹
شکل ۶-۴ تصویر اول خروجی الگوریتم (تصویر شماره صفر).....	۲۹
شکل ۷-۴ تصویر دوم خروجی الگوریتم (تصویر شماره یک).....	۳۰
شکل ۸-۴ تصویر سوم خروجی الگوریتم (تصویر شماره دو).....	۳۰
شکل ۹-۴ تصویر چهارم خروجی الگوریتم (تصویر شماره سه).....	۳۰
شکل ۱۰-۴ تصویر پنجم خروجی الگوریتم (تصویر شماره پنج).....	۳۱
شکل ۱۱-۴ تصویر ششم خروجی الگوریتم (تصویر شماره شش).....	۳۱

فهرست علائم

علائم لاتین

p	مرکز دایره
n	بردار نرمال
r	شعاع دایره
c	مرکز کره

زیرنویس‌ها

i	شماره فریم عکس‌ها
---	-------------------

فصل اول

مقدمه

چشم‌ها دریچه ارتباط ما با دنیای بیرون هستند و اگر بتوانیم مسیر نگاه یک فرد را دنبال کنیم، می‌توانیم به مواردی که مورد توجه وی است پی ببریم و نحوه دریافت او از محیط اطراف را مورد ارزیابی قرار دهیم [1]. ردیابی چشم دارای کاربردهای بسیاری در حوزه‌های مختلف از جمله تکنیک‌های تعامل انسان و رایانه، تشخیص پزشکی، مطالعات روان‌شناختی و دید کامپیوتری است [2].

زاویه سمت نگاه یک مولفه قابل‌مشاهده بیرونی است که توجه بصری انسان را نشان می‌دهد و بسیاری تلاش کرده‌اند آن را ثبت کنند. مفهوم ردیابی سمت نگاه به اواخر قرن هجدهم باز می‌گردد. امروزه راه‌حل‌های مختلفی برای این کار وجود دارد که بسیاری از آن‌ها تجاری هستند، اما همه آن‌ها دارای یک یا چند مورد از مشکلات هزینه بالا، سخت‌افزار آسیب‌زا و یا کاهش دقت در شرایط بلادرنگ هستند. این عوامل مانع از تبدیل ردیابی چشم به یک فناوری فراگیر می‌شوند که باید برای هر کسی با یک دوربین معقول در دسترس باشد. در این پروژه هدف ما غلبه بر این چالش‌ها است.

در مطالعات پیشین این حوزه، الگوریتم‌هایی مانند روش‌های مدلسازی [3] و روش‌های مبتنی بر یافتن همبستگی بین متغیرها [2] ارائه شده‌است. هرچند اکثر این روش‌ها از تصاویر باکیفیت بسیار بالا و تهیه شده تحت شرایط خاص استفاده می‌کنند. تصاویر ورودی این الگوریتم‌ها توسط دستگاه‌های تصویربرداری تشدید مغناطیسی و یا دوربین‌های خاص تهیه شده‌اند. این امر استفاده از این تکنولوژی به صورت عملی را دشوار می‌سازد.

ما در این پروژه از الگوریتم مدل‌سازی ارائه شده در پژوهش‌های قبلی [3] استفاده خواهیم کرد. با توجه به اینکه این الگوریتم برای عکس‌های تهیه شده بصورت عادی قابل‌استفاده نیست، از یک مرحله ناحیه‌بندی مردمک چشم [4] در تصاویر قبل از مدل‌سازی استفاده خواهیم کرد. با استفاده از این روش جزئیات نامرتبط از تصاویر ورودی الگوریتم مدل‌سازی حذف می‌شوند.

جهت تهیه تصاویر از چشم، سخت‌افزاری مشابه عینک طراحی می‌شود. این سخت‌افزار باید برای همه افراد قابل استفاده باشد. همچنین استفاده از آن مانع دید فرد نشود تا بتوانیم از آن در تحقیقات آتی و معاینات چشم فرد استفاده کنیم. با توجه به محدود بودن الگوریتم‌های مورد بررسی به چشم فرد، سخت‌افزار باید به گونه‌ای طراحی شود که تنها از چشم فرد تصویر تهیه کند و صورت فرد بطور کامل در تصویر قرار نگیرد.

در بخش اول این پایان نامه به بررسی مطالعات انجام شده پیشین می پردازیم و روش های ردیابی حرکات چشم را مرور می کنیم، سپس الگوریتم مورد استفاده در این پژوهش را مورد بررسی قرار داده و پس از طراحی سخت افزار مورد نیاز برای ردیابی حرکات چشم، پروژه را در فضای زبان برنامه نویسی پایتون پیاده سازی می کنیم.

فصل دوم

بررسی مطالعات انجام شده

بررسی مطالعات انجام شده

الگوریتم‌های ردیابی حرکات چشم انواع گوناگونی دارند که بسته به شرایط و کاربرد آن در پژوهش مورد استفاده قرار می‌گیرند. در این بخش به مرور برخی از این الگوریتم‌ها می‌پردازیم و پس از بررسی مقالات و مطالعات پیشین الگوریتم مناسب را انتخاب می‌کنیم.

۱-۲- انواع الگوریتم‌ها

الگوریتم‌های ردیابی حرکات چشم مبتنی بر دوربین معمولاً به دو دسته تقسیم می‌شوند: الگوریتم‌های تشخیص مردمک چشم [5] و الگوریتم‌های تشخیص سمت نگاه [3]. الگوریتم‌های تشخیص مردمک چشم در فضای دوبعدی عمل می‌کنند و تلاش می‌کنند تا مردمک چشم را بصورت یک بیضی تشخیص دهند. این الگوریتم‌ها از روش‌هایی شامل تشخیص تیره‌ترین نقاط تصویر به عنوان مردمک چشم [3] و روش‌های مبتنی بر یادگیری ماشین [5] جهت تشخیص مردمک چشم در تصاویر استفاده می‌کنند. الگوریتم‌های تشخیص سمت نگاه [3] از دو مرحله تشخیص مردمک چشم و تشخیص سمت نگاه تشکیل می‌شوند. این الگوریتم‌ها پس از تشخیص مردمک، اطلاعات بدست آمده از مردمک چشم در تصاویر را در فضای سه بعدی دنیای اطراف به زاویه سمت نگاه و یا نقطه مورد توجه چشم تبدیل می‌کنند. نقطه مورد توجه چشم جسمی در فضای اطراف است که فرد در لحظه به آن نگاه می‌کند.

۲-۲- الگوریتم‌های تشخیص سمت نگاه

بیشتر الگوریتم‌های تشخیص سمت نگاه می‌توانند به دو دسته الگوریتم‌های مبتنی بر همبستگی بین متغیرها و الگوریتم‌های مبتنی بر مدل تقسیم شوند. الگوریتم‌های مبتنی بر همبستگی بین متغیرها فرض می‌کنند یک رابطه نامعلوم بین متغیرهای بدست آمده از تصاویر چشم چشم و زاویه نگاه وجود دارد. و این رابطه را بطور تقریبی با استفاده از روش‌های همبستگی و معمولاً همبستگی چندجمله‌ای، بدست می‌آورند. همچنین برخی از این روش‌ها از شبکه‌های عصبی و روش‌های یادگیری عمیق برای بدست آوردن این رابطه استفاده می‌کنند.

الگوریتم‌های مبتنی بر مدل تلاش می‌کنند با استفاده از اطلاعات چشم ابتدا مدل چشم را در فضای سه بعدی تشکیل دهند و سپس زاویه نگاه را بدست آورند. ساخت مدل سه‌بعدی چشم شامل بدست آوردن مرکز کره چشم و شعاع کره چشم است. حرکات مردمک چشم را می‌توان بصورت حرکات دیسکی مماس بر روی کره چشم در نظر گرفت. در این حالت بردار نرمال مردمک چشم، یعنی خط بین مرکز دیسک مردمک و مرکز کره چشم در هر لحظه به عنوان بردار نگاه فرد اعلام می‌شود [3].

در این پروژه به طور خاص دو پژوهش قبلی جهت تشخیص مردمک و یک الگوریتم جهت مدل‌سازی کره چشم مورد بررسی قرار گرفت. در پژوهش اول [3] برای تشخیص مردمک چشم تیره‌ترین نقاط تصویر به عنوان مردمک انتخاب شده و سپس با مدل‌سازی زاویه نگاه محاسبه می‌شود. الگوریتم ارائه شده جهت مدل‌سازی جهت استفاده در این پروژه مناسب است، اما برای تشخیص مردمک به الگوریتم با دقت بالاتری نیازمندیم. زیرا این الگوریتم ممکن است در شرایطی که نور محیط به دقت تنظیم نشده باشد، مژه چشم و یا سایه ناشی از عینک را به عنوان مردمک چشم تشخیص دهد. این امر موجب می‌شود مرحله مدل‌سازی چشم دچار خطای بسیاری شود. در پژوهش دوم [5] از روش ناحیه‌بندی تصاویر جهت تشخیص مردمک استفاده شده بود. براساس آزمایش‌های انجام شده در این پژوهش، این روش برای استفاده در تصاویر بدون تنظیم نور محیط نیز مناسب است و حتی در شرایطی که چشم فرد نیمه بسته باشد و یا سایه‌های محیط در تصاویر وجود داشته باشند می‌تواند با دقت بالای ۹۰ درصد مردمک چشم را تشخیص دهد.

۲-۳- روش‌های واسنجی

روش‌های ذکر شده در بالا برای تشخیص سمت نگاه به واسنجی نیاز دارند. واسنجی می‌تواند برای پیدا کردن رابطه بین متغیرهای چشم و سمت نگاه و یا برای اختصاص یک مدل به چشم فرد برحسب متغیرهای بدست آمده باشد. معمولاً برای واسنجی از کاربر خواسته می‌شود که به چند نقطه در فاصله معلوم نگاه کند. برای مثال برخی از مطالعات از روش واسنجی ۹ نقطه‌ای استفاده می‌کنند. روش‌های مبتنی بر مدل معمولاً به جای انجام واسنجی برای هر فرد، از میانگین داده‌های انسانی برای کاهش تعداد متغیرها استفاده می‌کنند، با این حال آن‌ها باز هم به واسنجی نیاز دارند.

برخی از روش‌ها از نورهای تنظیم شده استفاده می‌کنند و در صورت استفاده از این روش تشخیص سمت نگاه در شرایطی که منبع نور کنترل شده نباشد دشوار است. برخی از روش‌ها نیز از چند دوربین استفاده می‌کنند [1] تا بتوانند مدل سه‌بعدی چشم را با اطلاعات از چند زاویه مختلف بدست آورند.

با این حال همیشه استفاده از تصاویر باکیفیت، کنترل شده و تحت نور محیطی تنظیم شده امکان‌پذیر نیست. تعداد کاربردهای الگوریتم‌های تشخیص سمت نگاه با استفاده از سخت افزارهای خانگی و وب‌کم نصب شده بر روی فریم عینک و استفاده از نور LED و یا مادون قرمز در حال افزایش است. در چنین شرایطی استفاده از روش‌های تنظیم کردن نور بسیار دشوار و ناممکن است.

۲-۴- جمع‌بندی فصل

در این فصل به بررسی انواع مختلف الگوریتم‌های ارائه شده پیشین برای کاربردهای ردیابی حرکات چشم پرداختیم. هدف ما در این پروژه تشخیص سمت نگاه است. در نتیجه از یک الگوریتم جهت تشخیص مردمک و از یک الگوریتم جهت تشخیص سمت نگاه استفاده خواهیم کرد. همچنین می‌خواهیم ردیابی سمت نگاه در محیطی عادی و بدون استفاده از نورهای محیطی تنظیم شده صورت گیرد، به همین علت به الگوریتمی با دقت بسیار بالا جهت تشخیص مردمک چشم نیازمندیم. جهت تشخیص مردمک از الگوریتمی مبتنی بر یادگیری عمیق استفاده می‌کنیم. الگوریتم‌های مبتنی بر یادگیری عمیق به علت داشتن دارای کاربردهای بسیاری در زمینه تصویربرداری و تشخیص پزشکی هستند [5]. برای تشخیص سمت نگاه از یک الگوریتم مبتنی بر مدل [3] استفاده خواهیم کرد. در بخش بعدی به شرح و جزئیات فنی الگوریتم‌های مورد استفاده در این پروژه خواهیم پرداخت.

فصل سوم

الگوریتم ردیابی سمت نگاه

مقدمه فصل

روش مورد استفاده در این تحقیق براساس مدل سازی کره چشم است. پس از تشخیص مردمک چشم به عنوان بیضی در تصویر، زاویه و محل قرارگیری کره چشم را در فضای سه بعدی تخمین می زنیم. با ترکیب اطلاعات بدست آمده از تصویرهای مختلف برحسب این فرض که حرکات مردمک چشم محدود شده است، مدل سه بعدی چشم را می سازیم.

۳-۱- فرض ها و محدودیت ها

در این روش ما زاویه بین بردار چشم و بردار نگاه را در نظر نمی گیریم. زاویه بردار نگاه برابر با زاویه خط متصل کننده بین مردمک چشم فرد و نقطه ای است که فرد به آن می نگرد. زاویه بردار چشم فرد زاویه خطی است که از اتصال نقطه مرکز مردمک چشم و مرکز کره چشم به دست می آید. زاویه بین بردار چشم و بردار نگاه برای هر فرد مقداری کم و ثابت است و می توانیم این زاویه را با استفاده از واسنجی یک نقطه ای بدست بیاوریم. در نتیجه از این الگوریتم نمی توانیم بطور مستقیم برای تشخیص نقطه مورد توجه استفاده کنیم.

همچنین ما حرکات سر را در نظر نمی گیریم. زیرا فرض می کنیم الگوریتم برای سخت افزاری مانند عینک مورد استفاده قرار می گیرد و در آن با حرکات سر رابطه چشم و دوربین تغییر نمی کند. در این الگوریتم مهم است که موقعیت دوربین نسبت به چشم ثابت باشد و به طوری تنظیم شده باشد که چشم فرد در تصویر قابل مشاهده باشد.

۳-۲- شرح الگوریتم

روش عملکرد این الگوریتم به شرح زیر است:

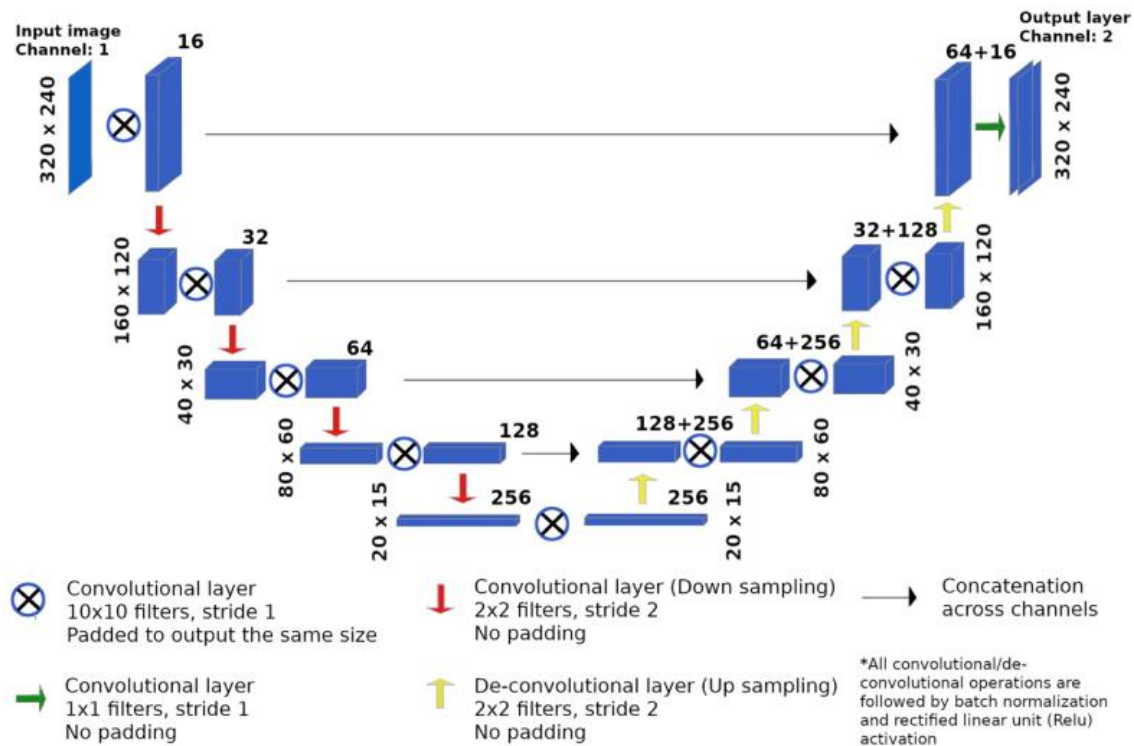
ابتدا در هر تصویر بصورت جداگانه موقعیت مردمک چشم را تشخیص می دهیم.

پس از بدست آوردن موقعیت مردمک چشم در هر تصویر، بصورت مستقل براساس هر مردمک بدست آمده یک دایره معادل با دایره مردمک چشم در فضای سه بعدی بدست می آوریم. و از ترکیب این دایره

های سه بعدی یک مدل برای کره چشم به دست می‌آوریم. مدل کره چشم شامل مرکز و شعاع کره چشم است. سپس این مدل را به نحوی بهینه‌سازی می‌کنیم که مردمک‌های بدست آمده به بهترین حالت ممکن بر روی مدل کره چشم منطبق شوند. مراحل این روش با جزئیات در بخش‌های بعدی این فصل شرح داده شده است.

۳-۲-۱ - بدست آوردن موقعیت مردمک چشم

مرحله اول بدست آوردن موقعیت مردمک چشم در هر فریم به صورت دوبعدی است. با توجه به تهیه تصاویر ورودی بدون تنظیم کردن نور در محیط آزمایشگاهی، به الگوریتمی با دقت بالا نیاز داریم. برای این کار از ناحیه‌بندی تصاویر به کمک الگوریتم CNN استفاده می‌کنیم. معماری‌های گوناگونی برای قطعه‌بندی در تصاویر پزشکی تاکنون ارائه شده‌اند. یکی از معماری‌های خوب برای کاربردهای پزشکی شبکه U-Net است، که برای حوزه‌های زیادی قابل استفاده است. ما از یک شبکه U-Net ساده برای جدا کردن مردمک در تصاویر تهیه شده استفاده می‌کنیم [3]. شبکه U-Net از لایه‌های مختلف برای استخراج ویژگی مورد نظر تشکیل شده است و شامل یک مسیر نمونه‌برداری کاهشی در سمت چپ و یک مسیر نمونه‌برداری افزایشی در سمت راست می‌باشد. برای حفظ ویژگی‌های تصاویر فرکانس بالا، مسیرهایی افقی از سمت نمونه‌برداری کاهشی به سمت نمونه‌برداری افزایشی وجود دارد. معماری استفاده شده به صورت زیر است.



شکل ۳-۱ معماری شبکه عصبی.

خروجی این شبکه شامل دو نقشه خروجی برای مردمک چشم و پس‌زمینه تصویر است. همچنین تصویر خروجی ابعادی برابر با تصویر ورودی دارد.

برای مرحله آموزش این شبکه عصبی از دو دیتاست در مرکز بیماری‌های تعادلی آلمان استفاده شده‌است. این دیتاست‌ها شامل تصاویری از چشم ۶۲ فرد سالم با میانگین سنی ۲۷ سال هستند و به کمک دستگاه MRI تهیه شده‌اند. از بین این دو دیتاست به‌صورت تصادفی ۳۹۴۶ فریم انتخاب شده و به عنوان داده‌های آموزش مورد استفاده قرار گرفتند. در مرحله تست، علاوه بر این دو دیتاست از پنج دیتاست جدید استفاده شده‌است و پنج نفر با علامت‌گذاری دستی محل مردمک چشم در تصاویر با استفاده از روش فاصله اقلیدسی دقت این الگوریتم را محاسبه کردند. دقت این الگوریتم ۹۳ درصد ارزیابی شده‌است.

ما برای این مرحله از سامانه از چهارچوب deepvov که بدین منظور ارائه شده‌است، و وزن‌های بدست آمده در پژوهش بالا استفاده کرده‌ایم [5].

۳-۲-۲- بدست آوردن مکان هندسی مردمک

مرحله دوم الگوریتم هر تصویر دوبعدی مردمک را به مردمک متناظر آن در فضای سه بعدی تبدیل می‌کند [6]. در واقع باید دایره‌ای در فضای سه بعدی بدست آوریم که تصویر آن در دوربین برابر با تصویر مردمک بدست آمده است. برای این کار می‌توانیم از دو روش تجسم ضعیف و یا تجسم قوی استفاده کنیم. تجسم ضعیف تنها تقریبی از تجسم قوی است و فقط برای اشیای دوری که در نزدیک محور دوربین هستند کاربرد دارد. اگر مردمک چشم نزدیک به دوربین باشد و یا از محور دوربین دور شود، این روش پاسخگو نخواهد بود. به همین دلیل ما از روش تجسم قوی استفاده می‌کنیم. این روش در ادامه شرح داده می‌شود.

برای بدست آوردن مردمک چشم در فضای سه بعدی با روش تجسم قوی مکان هندسی مردمک چشم می‌تواند بصورت مخروطی در نظر گرفته شود که قاعده آن مردمک چشم و راس آن نقطه کانونی لنز دوربین باشد. تصویر بیضی شکل مردمک که در مرحله قبل بدست آورده‌ایم، محل تقاطع این مخروط و صفحه تصویر است. با توجه به داشتن تقاطع صفحه تصویر و مخروط و نقطه کانونی دوربین می‌توانیم مردمک چشم در فضای سه بعدی را بدست آوریم.

در نتیجه برای بدست آوردن دایره مردمک می‌توانیم این مخروط را بازسازی کنیم. دایره‌های سه بعدی که می‌توانند در مخروط به نحوی قرار بگیرند که تصویر دوبعدی آن‌ها برابر با تصویر مردمک به دست آمده در مرحله قبل باشد، دایره‌های احتمالی برای مردمک چشم هستند. برای بازسازی مخروط و بدست آوردن دایره‌های مردمک از متد صفایی-راد [6] استفاده می‌کنیم.

با استفاده از این روش می‌توانیم موقعیت مردمک، زاویه نگاه و شعاع مردمک را برای یک تصویر بدست آوریم. زاویه نگاه برابر با بردار نرمال مردمک یعنی خط گذرا از مرکز دایره به دست آمده و عمود بر صفحه دایره مردمک است.

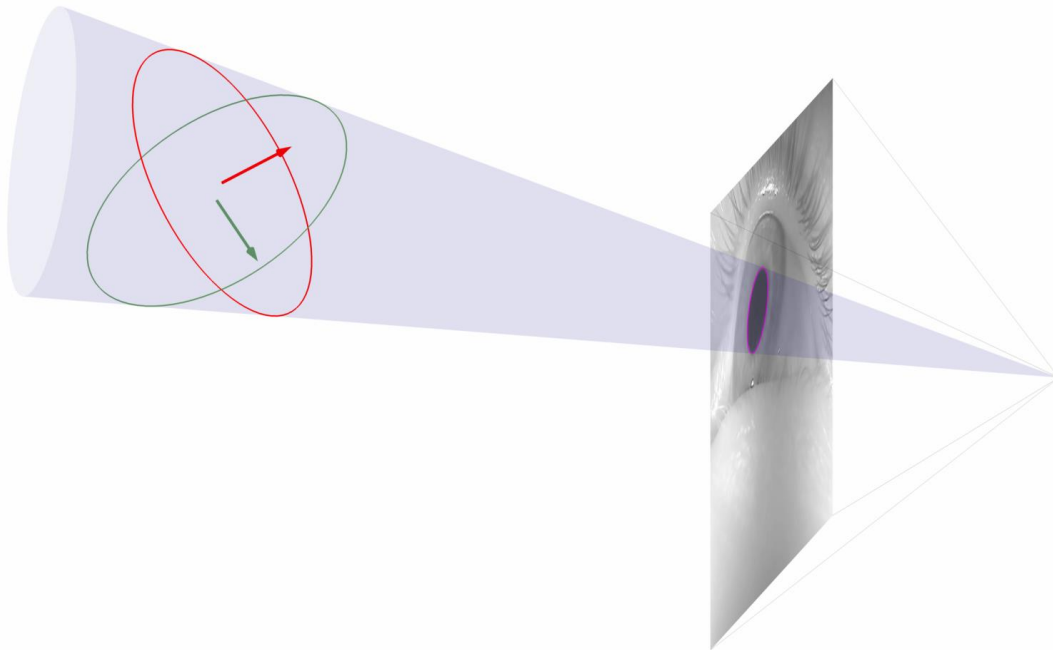
$$\text{pupil circle} = (p, n, r) \quad (1-3)$$

با توجه به اینکه دو بردار نرمال برای هر دایره وجود دارد، در صورتی که لازم باشد بردار نرمال مردمک ۱۸۰ درجه دوران پیدا می‌کند تا رو به دوربین و راس مخروط قرار گیرد.

در این بخش برای هر بیضی مردمک در فضای دوبعدی دو ابهام وجود دارد. ابهام اول مربوط به اندازه است برای تبدیل هر عکس دو بعدی به سه بعدی مشخص نیست که مردمک کوچک است و نزدیک به دوربین قرار گرفته و یا بزرگ است و در فاصله ی دور از دوربین قرار گرفته است. ما این مشکل را با قرار دادن مقدار دلخواه به جای شعاع دایره حل کردیم و در مراحل بعدی مقدار دقیق شعاع مردمک را بدست می آوریم. در واقع هر دایره که صفحه آن موازی یکی از پاسخ های بدست آمده باشد و بصورت مماس در درون مخروط قرار بگیرد می تواند پاسخ دیگر این مساله باشد.

ابهام دوم این است که همانطور که در شکل ۲-۳ نشان داده شده است، برای هر شعاع مردمک فرضی دو بیضی به عنوان راه حل بدست می آید. این بیضی ها نسبت به محور اصلی مخروط متقارن هستند و از حل یک معادله ی درجه ی چهار به دست می آیند. در این مرحله از الگوریتم، ما ابهام بین این دو حالت را در نظر نمی گیریم و هر دو راه حل را به صورت زیر باز می گردانیم:

$$(p+, n+, r) , (p-, n-, r) \quad (2-3)$$



شکل ۲-۳ بدست آوردن مردمک.

۳-۲-۳- بدست آوردن مدل چشم

برای بدست آوردن بردار نگاه، ما فقط جهت و محل مردمک را در نظر می‌گیریم و به مدل کامل چشم نیاز نداریم. زیرا فقط می‌خواهیم مردمک و محدوده‌ی حرکت آن روی کره چشم را مدل‌سازی کنیم و به پلک و جزئیات کره‌ی چشم نیازی نداریم.

در نتیجه به جای مدل‌سازی مردمک چشم به صورت یک حفره‌ی استوانه‌ای در کره‌ی چشم، ما آن را به صورت یک دیسک که مماس بر روی سطح یک کره‌ی چرخان قرار گرفته است مدل‌سازی می‌کنیم. این کره دارای مرکز دوران یکسانی با کره‌ی چشم می‌باشد. با صرف نظر از اختلاف زاویه بین بردار چشم و بردار نگاه، بردار نگاه برابر با بردار نرمال دیسک و یا یک بردار از مرکز کره به مرکز دایره مردمک است.

۳-۲-۴- تخمین مرکز کره چشم

حال به کمک اطلاعات به دست آمده در مرحله دوم، برای هر تصویر بدست آمده از چشم، دایره‌ی مردمک (p_i, n_i, r_i) ، که در آن i شماره‌ی فریم، p مرکز دایره، n بردار نرمال یا بردار سمت نگاه، و r شعاع دایره است را در نظر می‌گیریم. با توجه به این مدل از مردمک، می‌خواهیم کره‌ای را بدست آوریم که بر کل دایره‌های مردمک مماس است. با توجه به اینکه کل مردمک‌های بدست آمده بر کره چشم حقیقی مماس هستند، بردار نرمال دایره‌ها بردار شعاعی کره است و در نتیجه محل تقاطع آنها مرکز کره می‌باشد.

با این حال دو مشکل وجود دارد که مربوط به دو ابهام مطرح شده در قبل می‌باشد. مشکل اول این است که ما موقعیت دقیق سه‌بعدی هر دایره‌ی مردمک را نمی‌دانیم. تنها موقعیت دوایر با این فرض که مردمک دارای شعاع خاصی است را می‌دانیم. این مشکل ابهام اندازه-فاصله نامیده می‌شود. مردمک‌های دیگر که پاسخ این مساله هستند و شعاع دیگری دارند می‌توانند از تجانس این دایره به دست آیند. با توجه به این فرض اگر شعاع مردمک در طول تصویرها تغییر نکند و دچار انقباض یا انبساط نشود، موقعیت مربوط به آن با در نظر گرفتن مقیاس، درست خواهد بود. این حالت برای روش‌هایی درست است که از عنبیه به جای مردمک استفاده می‌کنند. هرچند به علت بزرگ شدن مردمک، ما نمی‌توانیم از این فرض استفاده کنیم.

مشکل دوم این است که برای هر بیضی دو دایره به جای یک دایره به دست می‌آید و ما نمی‌دانیم کدام یک از آن دو درست هستند. ما در این مرحله هر دو مشکل را با در نظر گرفتن تقاطع بردارهای نرمال n

در فضای دوبعدی به جای فضای سه‌بعدی حل می‌کنیم. در این مورد، ابهام اندازه-فاصله ناپدید می‌شود. زیرا از تصویر اصلی استفاده می‌کنیم.

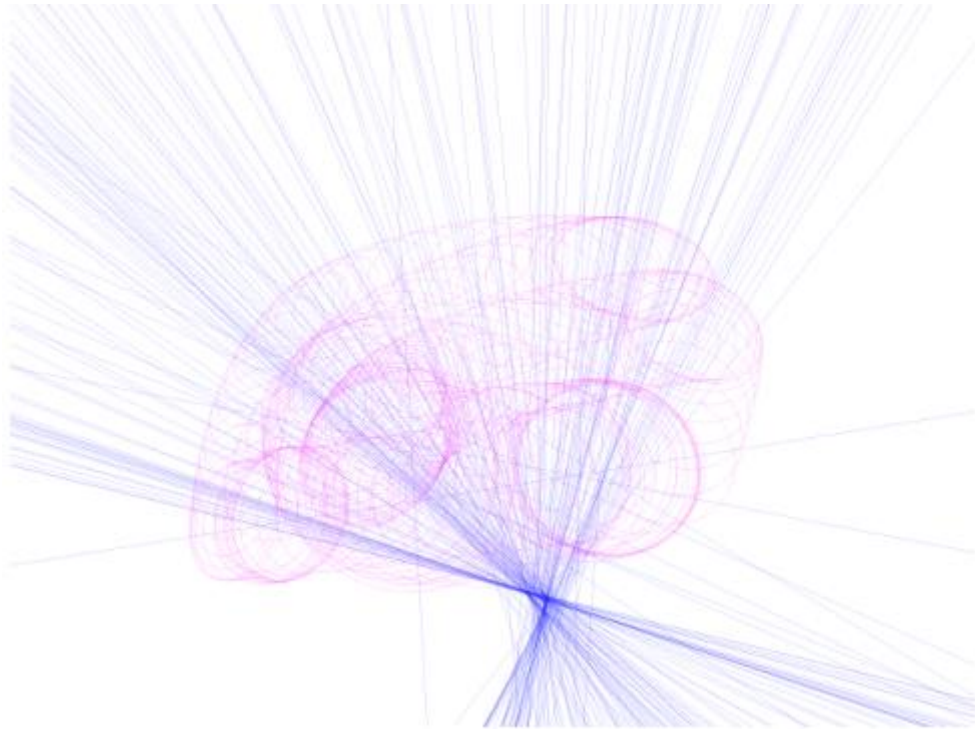
همچنین ابهام وجود دو دایره نیز از بین می‌رود، زیرا بردار نرمال هر دو دایره در فضای دوبعدی موازی هستند. به طور مشابه خط بین مرکز دو دایره‌ی تصویر شده و مرکز کره در فضای دوبعدی موازی با بردار \mathbf{n} می‌باشد. در نتیجه می‌توانیم یکی از دو دایره را به دلخواه برای این مرحله از الگوریتم انتخاب کنیم و انتخاب هر یک از آن‌ها تاثیری در نتیجه نهایی نخواهد داشت. پس از انتخاب دایره‌ها، می‌توانیم مرکز کره را با محاسبه کردن محل تقاطع خطوط نرمال دایره‌های انتخاب شده محاسبه کنیم. این خط‌ها متناظر با زاویه‌ی سمت نگاه هستند. هر خط از مرکز مردمک تصویر شده می‌گذرد و با بردار نرمال مردمک موازی است. محل تقاطع خطوط L را می‌یابیم:

$$L_i = \{(x, y) = \tilde{\mathbf{p}}_i + s\tilde{\mathbf{n}}_i \mid s \in \mathbb{R}\} \quad (3-3)$$

از آنجایی که ممکن است خط‌های اندازه‌گیری و عددی در این بردارها وجود داشته باشد، این خطوط در یک نقطه‌ی ثابت تقاطع نخواهند داشت. به جای آن ما نقطه‌ای که به نقاط تقاطع خطوط نزدیک‌تر است را به کمک روش کمترین مجموع مربعات به دست می‌آوریم.

$$\tilde{\mathbf{c}} = \left(\sum_i I - \tilde{\mathbf{n}}_i \tilde{\mathbf{n}}_i^T \right)^{-1} \left(\sum_i (I - \tilde{\mathbf{n}}_i \tilde{\mathbf{n}}_i^T) \tilde{\mathbf{p}}_i \right) \quad (4-3)$$

نقطه بدست آمده مرکز کره چشم است. سپس در مرحله بعد به کمک مرکز کره‌ی به دست آمده، تصویر سه‌بعدی کره را به دست می‌آوریم.



شکل ۳-۳ بدست آوردن مرکز کره.

۳-۲-۵- تخمین شعاع کره

پس از اینکه تصویر مرکز کره را به دست آوردیم باید توجه کنیم که بردار نرمال هر مردمک باید به سمت خارج کره اشاره کند، در نتیجه رابطه زیر برقرار است.

$$n_i \cdot (c - p_i) > 0 \quad (۵-۳)$$

بردار نرمال مردمک تصویر شده نیز باید به سمت خارج مرکز کره تصویر شده اشاره کند. در نتیجه داریم:

$$\tilde{n}_i \cdot (\tilde{c} - \tilde{p}_i) > 0 \quad (۶-۳)$$

از آنجایی که بردارها در دو دایره بدست آمده به ازای هر تصویر و شعاع خاص مقارن هستند، بردارهای دو دایره به دو جهت مخالف هم اشاره می‌کنند. به این معنی که یکی از آن‌ها به سمت داخل کره و دیگری به سمت خارج کره اشاره می‌کند. در این حالت می‌توانیم ابهام وجود دو دایره را حل کنیم و دایره‌ای که بردار نرمال آن به سمت خارج کره اشاره می‌کند را انتخاب کنیم.

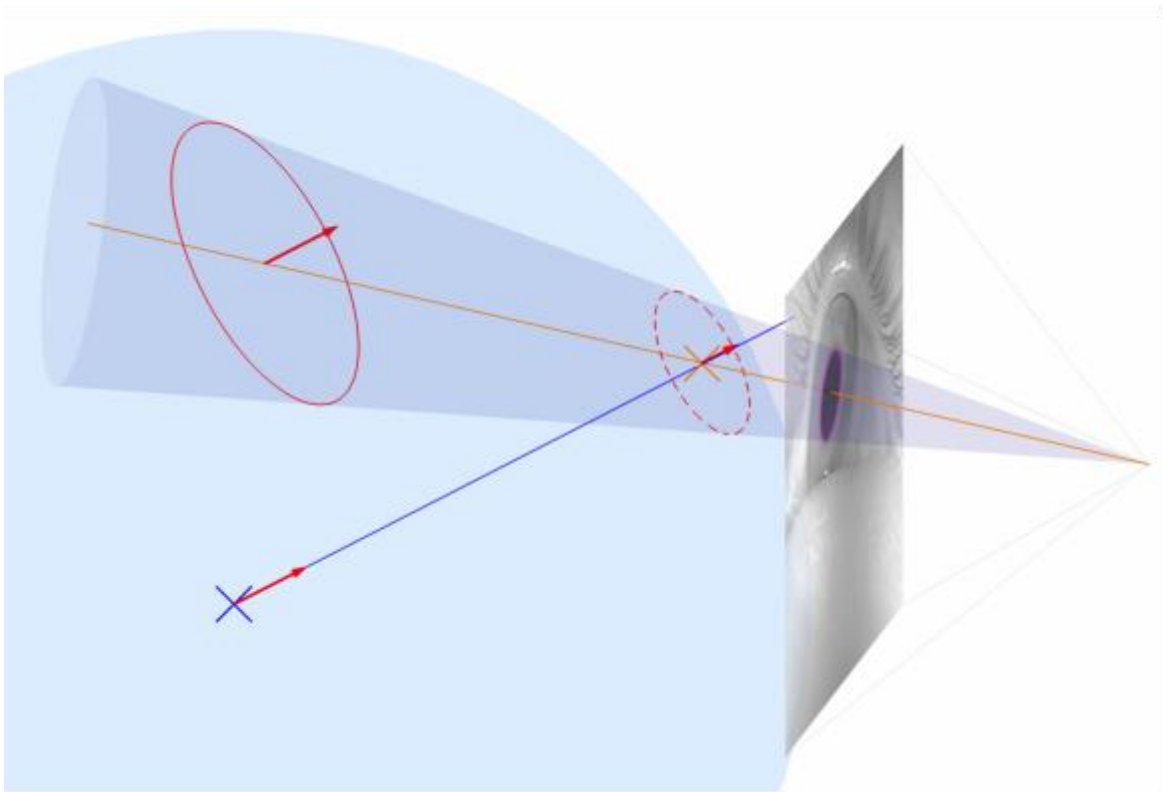
سپس می‌توانیم از تصویر مردمک‌های انتخاب شده برای تخمین شعاع کره استفاده کنیم. از آنجایی که هر مردمک روی کره قرار دارد، باید فاصله‌ی بین مرکز مردمک و مرکز کره، شعاع کره باشد.

هرچند بار دیگر به خاطر ابهام اندازه-فاصله در تصویرسازی مردمک و امکان تغییر اندازه‌ی واقعی مردمک نمی‌توانیم مرکز مردمک را به‌طور مستقیم به‌دست آوریم. به جای آن ما یک مرکز مردمک دیگر را که می‌تواند تصویر احتمالی مردمک باشد ولی در فاصله‌ی متفاوتی قرار دارد را در مرحله قبل انتخاب کرده بودیم. این به این معنی است که این نقطه باید جایی در بین خط احتمالی مرکز مردمک و مرکز دوربین قرار گرفته باشد. ما می‌خواهیم این دایره با فرضیات ما هماهنگ باشد. یعنی دایره مماس بر کره‌ی چشم قرار گرفته باشد. این به این معنی است که می‌خواهیم بردار نرمال آن موازی با خط گذرا از مرکز کره تا مرکز مردمک باشد.

با در نظر گرفتن محدودیت‌های ذکر شده در بالا، مرکز مردمک می‌تواند از تقاطع خط گذرا از مرکز کره به مرکز مردمک قبلی، و خط گذرا از مرکز مردمک قبلی و مرکز دوربین به دست آید. از آنجایی که خطوط در فضای سه بعدی لزوماً متقاطع نیستند، دوباره با استفاده از روش کمترین مجموع مربعات، نقطه‌ی احتمالی تقاطع را می‌یابیم.

شعاع کره از پیدا کردن میانگین فواصل از مرکز کره و مرکز مردمک‌ها به دست می‌آید.

$$R = \text{mean}(\{ R_i = \| \hat{p}_i - c \| \mid \forall i \}) \quad (7-3)$$



شکل 3-4 تخمین شعاع کره چشم.

۳-۲-۶- تخمین مردمک پایدار

با توجه به خطاهای احتمالی اندازه‌گیری، کل مردمک‌های به‌دست آمده در مرحله قبل روی کره قرار نمی‌گیرند. در این مرحله می‌خواهیم دایره‌های متناظر با مردمک‌ها مماس بر سطح کره‌ی به‌دست آمده در مرحله قبل باشند. در نتیجه می‌خواهیم مرکز هر مردمک روی سطح کره باشد. برای حل این مشکل مردمک‌های جدید را به نحوی به‌دست می‌آوریم که همه آن‌ها روی کره قرار گیرند. بدین منظور برای دایره‌های متناظر با هر مردمک باید روابط زیر برقرار باشند:

$$p'_i = sp_i \quad (۸-۳)$$

$$p'_i = c + Rn'_i \quad (۹-۳)$$

برای به دست آوردن مرکز مردمک جدید، باید مقدار s را طوری بیابیم که حاصل ضرب آن در مرکز مردمک قبلی بر روی سطح کره قرار گیرد. این مقدار می تواند با تقاطع خط و کره به دست آید.

این راه حل، دو جواب به ما می دهد که ما نزدیک ترین آن ها را انتخاب می کنیم، سپس بردار نرمال و شعاع مردمک جدید محاسبه می شود. توجه شود که به عنوان بخشی از این عملیات، بردار نرمال قبلی را پاک می کنیم. هنگامی که این مراحل انجام گرفت، یک مدل حرکات مردمک بدست می آید که در آن، هر دایره ی مردمک به صورت مماس بر روی کره ی ذکر شده، قرار گرفته است.

۳-۳- جمع بندی فصل

در این فصل الگوریتم مورد استفاده در پروژه مورد بررسی قرار گرفت. این الگوریتم ترکیبی از دو الگوریتم ناحیه بندی تصاویر و الگوریتم مدل سازی چشم است. با توجه به تهیه تصاویر و ورودی الگوریتم با نور تنظیم نشده بخش اول الگوریتم با دقت مناسبی مردمک چشم را بصورت دوبعدی تشخیص می دهد و سپس در ادامه با توجه به مردمک های تشخیص داده شده مدل کره چشم ساخته می شود. در این مدل مردمک چشم در هر تصویر تهیه شده از چشم بصورت یک دایره مماس بر کره چشم در نظر گرفته می شود. همچنین بردار نرمال مردمک یا همان شعاع کره که متصل کننده مرکز مردمک و مرکز کره است، بردار نگاه می باشد.

فصل چهارم پیاده‌سازی الگوریتم

پیاده‌سازی الگوریتم

در این بخش به پیاده‌سازی پروژه از نظر سخت‌افزاری و نرم‌افزاری می‌پردازیم. ابتدا سخت‌افزاری جهت فیلم‌برداری از چشم تهیه کرده و سپس بخش نرم‌افزاری پروژه را پیاده‌سازی می‌کنیم.

۴-۱- پیاده‌سازی بخش سخت‌افزاری

ردیاب چشم یک وسیله پوشیدنی برای کاربر و معمولاً به شکل دوربین ردیابی چشم یا عینک است. این نوع سیستم به طور معمول به یک دوربین یا یک آینه نیاز دارد تا در مسیر دیداری یک چشم یا هر دو قرار گیرد. ردیابی نگاه روی یک سیستم نصب‌شده روی سر نسبت به کل میدان دید انجام می‌شود و آن را برای آزمایش‌های واقعی ایده‌آل می‌سازد.

یک ردیاب چشم، بخش اصلی سخت‌افزار ردیابی چشم است که از نور مادون قرمز برای ایجاد روشنایی جهت ردیابی نگاه کاربر استفاده می‌کند. الگوریتم این پروژه برای سخت‌افزاری طراحی شده است که در آن یک یا چند دوربین بر روی یک عینک نصب شده باشند و موقعیت دوربین‌ها نسبت به چشم فرد ثابت باشد. در این حالت دوربین می‌تواند تصاویری با فاصله نزدیک از چشم فرد تهیه کند.

۴-۱-۱- محدودیت‌ها

در این بخش به بررسی محدودیت‌های این نوع سخت‌افزار و راه‌حل‌های پیش‌بینی شده برای آن‌ها می‌پردازیم.

۴-۱-۱-۱- نور خورشید

مانند تمام دستگاه‌های ردیابی چشم، این دستگاه نیز می‌تواند مشکلاتی برای ردیابی چشم در نور خورشید داشته باشد. بهترین سیستم‌های عینک، سیستم‌های مبتنی بر پوشش مادون قرمز هستند که می‌توانند به دستگاه متصل شوند. همچنین ردیابی چشم در شرایطی که کاربر به خاطر روشنایی و یا درخشش بیش از حد محیط چشم‌هایش را نیمه‌بسته کند، مشکل خواهد بود. در این حالت، یک محافظ مادون قرمز لازم است و ممکن است ایجاد سایه برای چشم کاربر با استفاده یک کلاه نیز مفید باشد.

۴-۱-۱-۲- حرکات غیرعادی چشم

از آنجا که دوربین‌های ردیابی چشم باید دید کامل و بدون مانع از چشم‌ها داشته باشند، ردیابی حرکت چشم به اطراف می‌تواند مشکل باشد و اغلب دقت کمتری دارد. همچنین گاهی اوقات ردیابی حرکات چشم، زمانی که شرکت‌کننده به یک شی در دستش مانند تلفن همراه نگاه می‌کند می‌تواند چالش برانگیز باشد. برای برطرف کردن این مشکل می‌توانیم از چند دوربین با زوایای دید متفاوت استفاده کنیم و میانگین نتایج مربوط به آن‌ها را به عنوان زاویه سمت نگاه در نظر بگیریم.

۴-۱-۱-۳- سیستم مختصات نسبی

برخلاف انواع دیگر سخت‌افزارهای مورد استفاده برای ردیابی چشم، هیچ سیستم مختصات مطلق در هنگام استفاده از یک دستگاه ردیابی چشم متحرک وجود ندارد. سیستم داده‌ها را در یک سیستم مختصات تعریف‌شده توسط دوربین صحنه ضبط می‌کند. این سیستم مختصات مانند صفحه فرضی عمل می‌کند که با سر شرکت‌کننده حرکت می‌کند. برای مثال، در یک سیستم از راه دور شما می‌توانید یک هدف متحرک را روی صفحه نمایش کامپیوتر نشان دهید. اگر موقعیت هدف را نسبت به صفحه را بدانید، تعیین اینکه آیا موقعیت نگاه هر شرکت‌کننده (که در پیکسل‌های صفحه قالب‌بندی شده است) با این هدف منطبق است ساده است. با این حال، با یک سیستم ردیابی چشم همراه، هدف ممکن است یک شی واقعی باشد که توسط دوربین ضبط شده است، مانند یک توپ فوتبال. موقعیت توپ در دوربین به موقعیت سر شرکت‌کننده بستگی دارد و می‌تواند به صورت تابعی از حرکت و حرکت توپ در یک زمان تغییر کند. و این برای هر کاربر متفاوت خواهد بود. این آنالیز نیاز به تفسیر ریاضی و ذهنی بیشتری دارد. با توجه به تغییرات چشمگیری که بین تجربه هر یک از شرکت‌کنندگان وجود دارد. هر یک از بازیکنان فوتبال به شکل متفاوتی در ویدئو حضور خواهند داشت و تجربه هر یک از کاربران باید به تنهایی مورد تحلیل قرار گیرد.

۴-۱-۲- دوربین‌ها

در این پژوهش از دو دوربین همراه AN99 برای ضبط تصاویر چشم استفاده شده است. مشخصات این دوربین‌ها در ادامه ذکر می‌شود. با توجه به وزن و نرخ تصویر بر ثانیه، این دوربین‌ها گزینه مناسبی برای

ردیابی حرکت چشم محسوب می‌شوند. این دوربین ها بصورت عمودی در بالای فریم عینک متصل می‌شوند.

- High resolution with 1/9 CMOS camera.
- Built-in 6 LED light (with Brightness and take Photo Control)
- Waterproof wire endoscope
- With a 5.5mm camera head.
- Photo shooting and video recording.
- Camera head outer diameter: 5.5mm.
- Pixels: 300000 pixels ON phone,1300000 pixels ON Computer
- Resolution: 640 x 480, 1280 x 480 (only on Computer) 30 fps
- LED: 6 White LED Bright light.
- Exposure Light: Automatic
- Focal distance: 5cm-infinite.
- Waterproof level: IP67.
- Support systems: Android /Windows 2000 / XP / Vista / 7 / 10. (Not supported for iOS)
- Temperature: 0 to 70 Celsius degree
- Photo Format: JPEG
- Video Format: AVI



شکل ۴-۱ دوربین مورد استفاده.

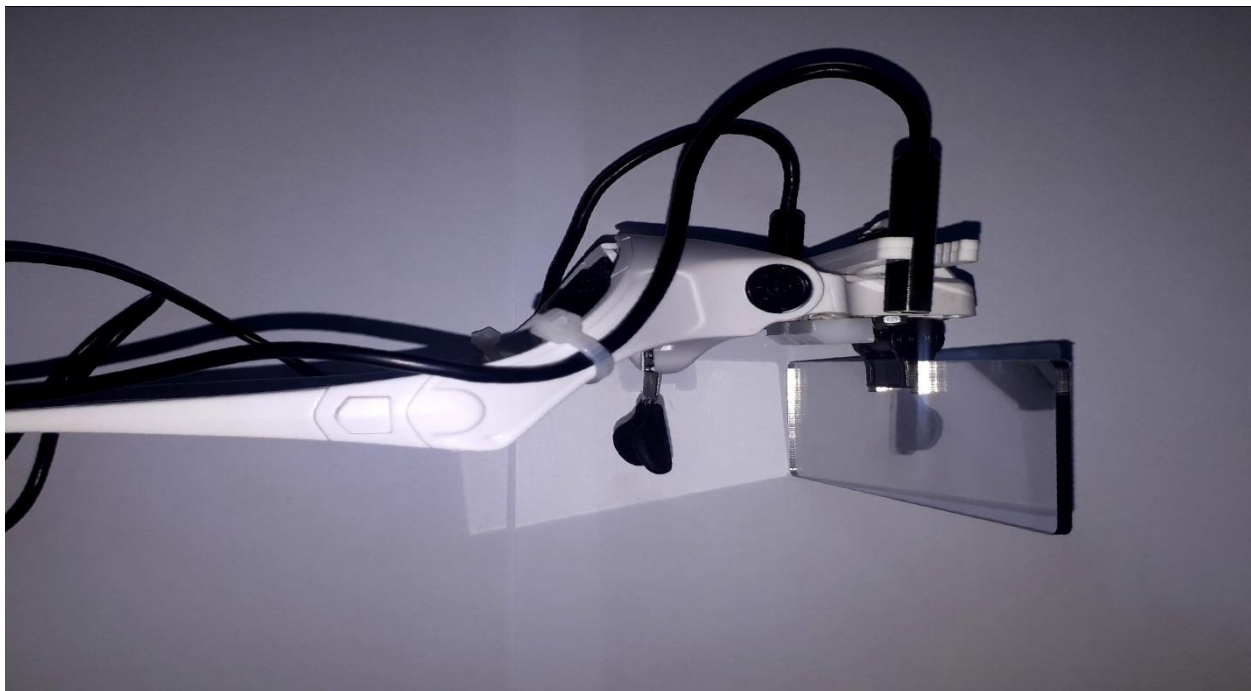
۴-۱-۳- ساختار سخت‌افزار

این سخت‌افزار از دو دوربین که در بالا شرح داده شد استفاده می‌کند. با توجه به اینکه در الگوریتم محاسبه نقطه توجه لحاظ نشده است نیازی به دوربین سوم برای فیلم‌برداری از محیط بیرونی نداریم. برای تصویر برداری از چشم از یک شیشه با روکش نیمه آینه‌ای استفاده شده است. بدین ترتیب با تنظیم زاویه مناسب شیشه دوربین می‌تواند از انعکاس چشم درون شیشه فیلم برداری کرده و مانعی برای دید فرد ایجاد نمی‌کند.

همچنین دوربین‌ها می‌توانند در بالای عینک بصورت افقی حرکت کنند تا فاصله دو دوربین برابر با فاصله بین دو چشم هر فرد تنظیم شود. و هر دوربین دقیقاً از مردمک یک چشم فیلم‌برداری کند.

برای تنظیم نور از یک منبع نور مادون قرمز و یک مقاومت استفاده شده است و این منبع نور از سه باتری جاسازی شده در عینک تغذیه می‌کند. استفاده از نور مادون قرمز به علت عدم تشکیل سایه باعث می‌شود الگوریتم دقت و عملکرد بهتری داشته باشد همچنین در صورت استفاده از یک فیلتر نور در مقابل دوربین و یا استفاده از دوربین مادون قرمز می‌توانیم از تغییرات نور ناخواسته جلوگیری کنیم. در این آزمایش به علت در دسترس نبودن دوربین مادون قرمز برای بالا بردن دقت الگوریتم، تصاویر در محیط با نور کم ضبط شدند.

برای انتقال تصاویر ضبط شده به کامپیوتر از پورت usb استفاده شد و تصاویر بصورت بلادرنگ به کامپیوتر منتقل می‌شوند. تصاویر استفاده شده در مراحل بعدی توسط این دوربین و عینک ضبط شده‌اند.



شکل ۴-۲ سخت‌افزار طراحی شده.

۴-۱-۴- مشکلات سخت‌افزار

استفاده از نور مادون قرمز موجب می‌شود استفاده طولانی مدت از این عینک برای چشم آزاردهنده باشد، هر چند در کاربردهایی مانند آزمایشات پزشکی که هدف اصلی این پروژه بودند و مدت زمان استفاده از سخت‌افزار در آن‌ها محدود است، استفاده از نور مادون قرمز بدون مشکل است. همچنین در صورتی که دوربین‌ها بطور مناسب تنظیم نشوند، الگوریتم با خطای زیادی برای تشخیص مردمک چشم مواجه خواهد شد، در چنین شرایطی ممکن است مژه‌های چشم و... به اشتباه مردمک تلقی شوند.

۲-۴- پیاده‌سازی بخش نرم‌افزاری

در این بخش به پیاده‌سازی الگوریتم و نرم‌افزار پروژه می‌پردازیم. کد این پروژه در زبان پایتون پیاده‌سازی شده است و بردار سمت نگاه را مطابق توضیحات ارائه شده در فصل قبل به دست می‌آورد.

۴-۲-۱- کتابخانه‌های مورد استفاده

در پیاده‌سازی این پروژه از چهارچوب deepvlog و کتابخانه‌های زیر استفاده شده است.

- numpy
- scikit-video
- scikit-image
- tensorflow-gpu
- keras

۴-۲-۲- واسنجی

در این پروژه از واسنجی استفاده نشده است. در صورتی که بخواهیم دقت بردار سمت نگاه بدست آمده را ارزیابی کنیم، ابتدا باید به کمک یک مرحله واسنجی زاویه بین بردار نگاه و بردار چشم فرد را به دست آوریم و سپس با اطلاع از این مقدار، مقادیر بردار نگاه را با مقادیر بردار چشم بدست آمده از الگوریتم مقایسه کنیم.

همچنین جهت بالا بردن دقت الگوریتم می‌توانیم از تنظیم نورهای محیط استفاده کنیم.

۴-۲-۳- پیاده‌سازی کد

در مرحله اول پیاده‌سازی از یک شبکه U-Net که به کمک کتابخانه keras پیاده‌سازی می‌شود و وزن‌های شرح داده شده در فصل قبل استفاده می‌کنیم. سودوکد این بخش بصورت زیر است:

جدول ۴-۱: شرح کد قطعه بندی.

01	program U-Net;
02	Get input images
03	Convert input to grayscale and proper shape

04	Build model
04	Define encoding stream
05	Define decoding stream
06	Define output layer activation function and operations
07	Load weights

سپس از تصاویر خروجی این مرحله به عنوان ورودی بخش بعد استفاده می‌شود. مراحل الگوریتم‌های تشخیص سمت نگاه بصورت کامل در بخش قبل شرح داده شده‌است.

۴-۲-۴- رابط کاربری

جهت سهولت استفاده از این برنامه برای کاربران یک رابط کاربری ساده به کمک flask پیاده‌سازی شده است، که در آن کاربران می‌توانند ویدیوهای خود را آپلود کرده و خروجی شامل فایل اکسل بردار نگاه در هر لحظه و ویدیوی خروجی را دریافت کنند.

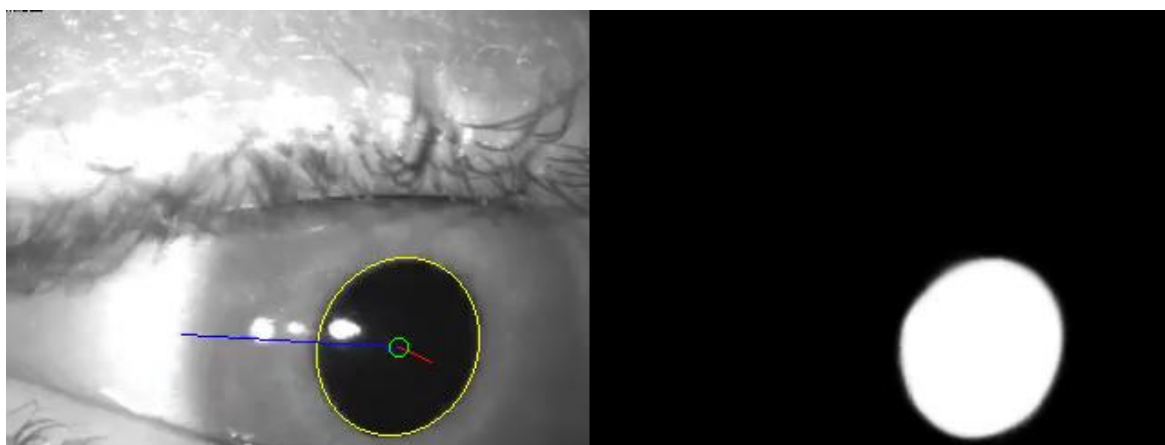
۴-۲-۵- آزمایش عملکرد پروژه

این پروژه را در دو حالت بدون استفاده از مرحله ناحیه‌بندی مردمک و با استفاده از ناحیه‌بندی مردمک تست کردیم. در حالت اول تنها الگوریتم در شرایط نور تنظیم شده قادر به تشخیص مردمک بود و در سایر موارد نقاط تیره‌تر تصویر شامل مژه‌ها و سایه‌ها را مردمک تشخیص می‌داد. اما در حالتی که از ناحیه‌بندی استفاده شد الگوریتم قادر به تشخیص مردمک در تصاویر تهیه شده بصورت عادی بود. در نتیجه در ادامه تنها از حالت استفاده از قطعه‌سازی مردمک استفاده می‌کنیم.

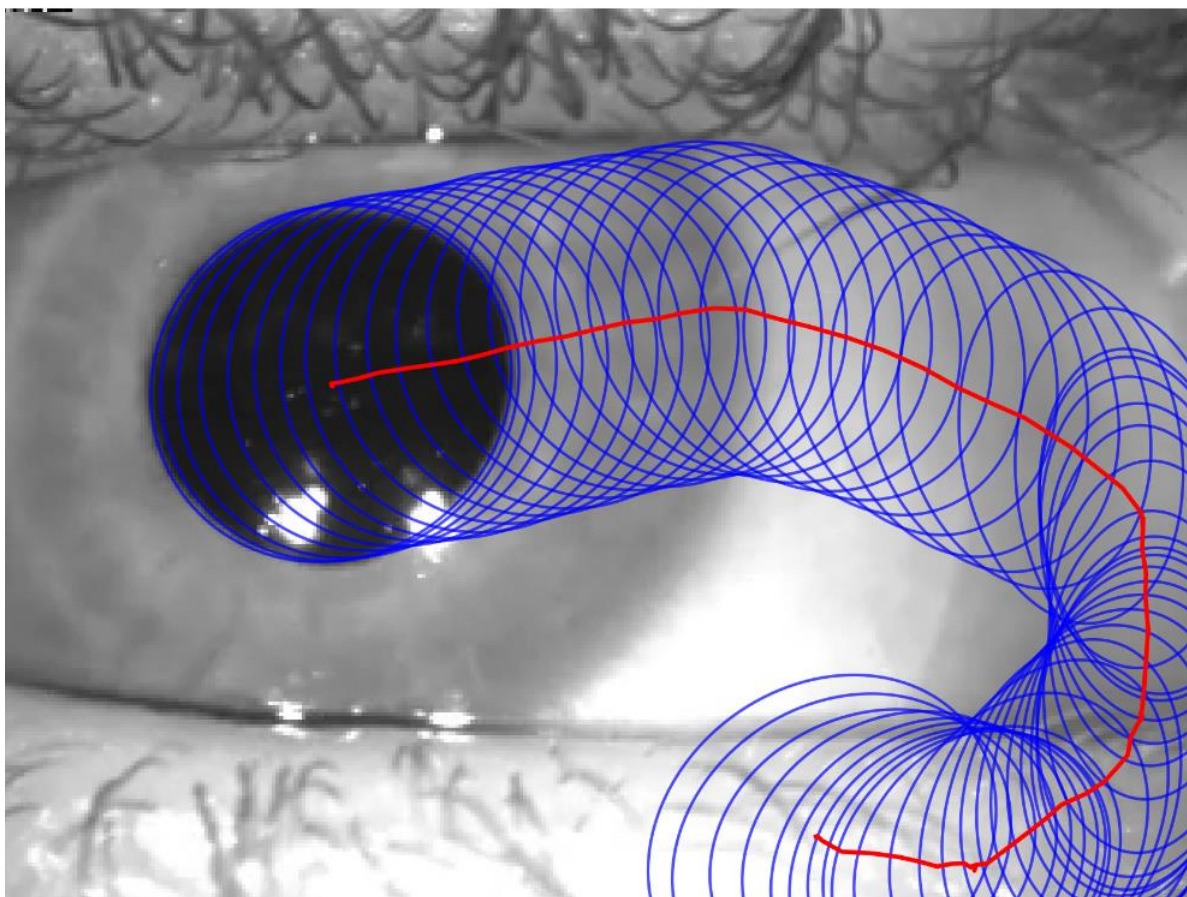
جهت بررسی دقیق‌تر عملکرد، از تصاویر تهیه شده به کمک سخت‌افزار پروژه به عنوان ورودی در نرم افزار تهیه شده در یکی از پژوهش‌های قبلی [3] استفاده کردیم. این نرم افزار نیز قادر به تشخیص مردمک در تصاویر تهیه شده نبود، و به همین علت نمی‌توانست مدل مردمک را بدست آورد.

۴-۳- نتایج بدست آمده

در این بخش نتایج بدست آمده توسط الگوریتم را بررسی می‌کنیم. ابتدا تصاویر ورودی بصورت دو بعدی پردازش می‌شوند. در زیر نمونه‌ای از خروجی بخش ناحیه‌بندی و مردمک‌های متوالی تشخیص داده شده در چند فریم ورودی را مشاهده می‌کنیم.



شکل ۳-۴ خروجی الگوریتم ناحیه‌بندی مردمک.



شکل ۴-۴ حرکت مردمک در تصاویر متوالی.

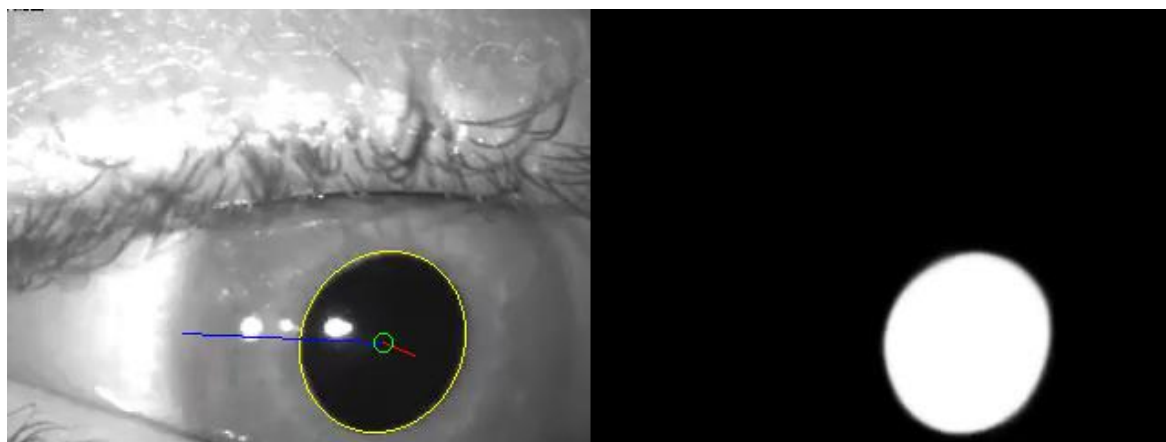
سپس مدل چشم ساخته شده و بردار سمت نگاه بدست می‌آید. خروجی الگوریتم به صورت فایل اکسل و ویدیو می‌باشد.

برای بررسی دقیق تر عملکرد کد و سنجش دقت آن اطلاعات شش لحظه را از فایل اکسل با تصاویر مربوط به آن لحظات مقایسه می‌کنیم. این تصاویر نمایشگر حالتی هستند که در آن مردمک چشم از سمت چپ به سمت راست حرکت می‌کند و در هر تصویر محل قرار گیری مردمک چشم و بردار نگاه را ثبت می‌کنیم.

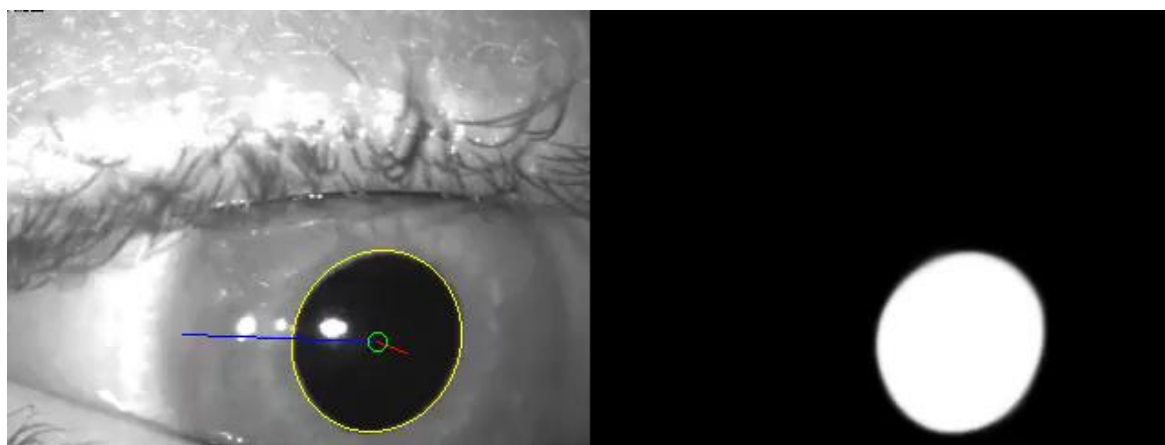
frame	pupil2D_x	pupil2D_y	gaze_x	gaze_y	confidence
0	202.864942	180.95245	71.45751	82.13459	0.983929
1	206.248397	181.9984	70.605605	81.78099	0.982818
2	214.544617	184.20783	68.414499	80.96978	0.980371
3	224.537129	185.25142	65.685389	80.30258	0.979309
4	233.120873	187.17658	63.167136	79.37378	0.980203
5	237.812493	189.451	61.64671	78.46794	0.979953

شکل ۴-۵ نتایج الگوریتم.

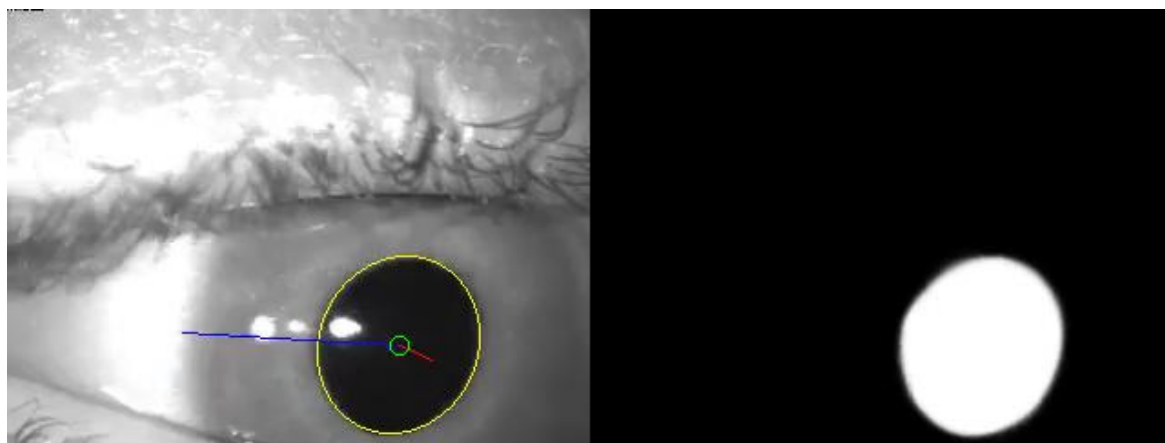
تصاویر مربوط به این مقادیر به ترتیب بصورت زیر می‌باشند.



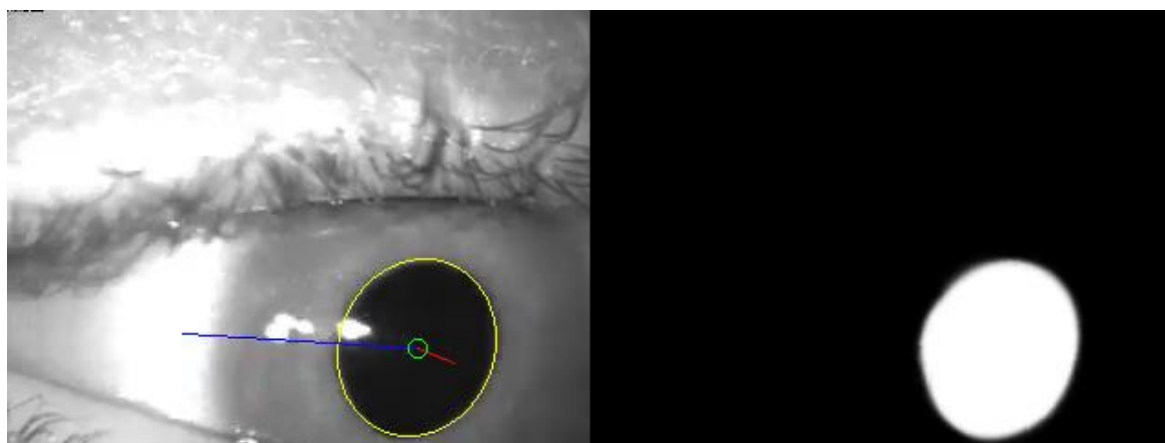
شکل ۴-۶ تصویر اول خروجی الگوریتم (تصویر شماره صفر).



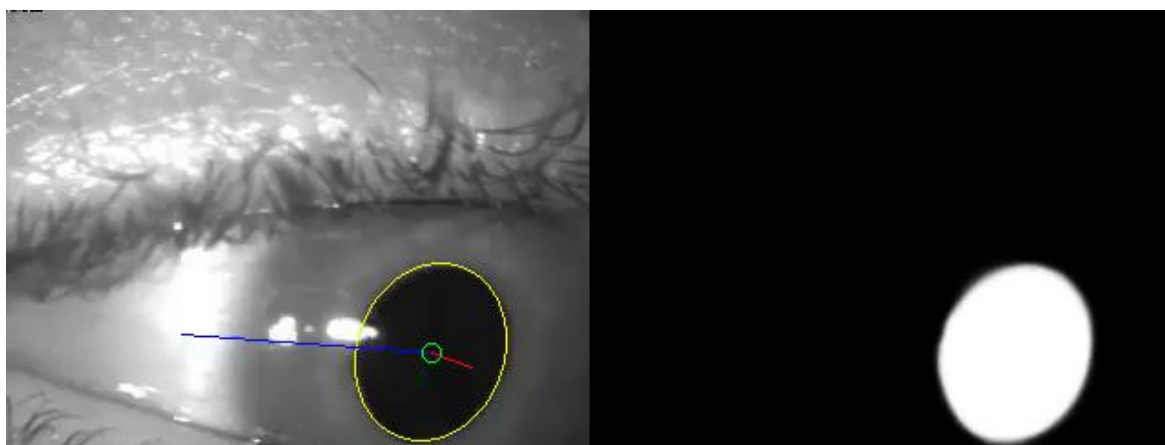
شکل ۴-۷ تصویر دوم خروجی الگوریتم (تصویر شماره یک).



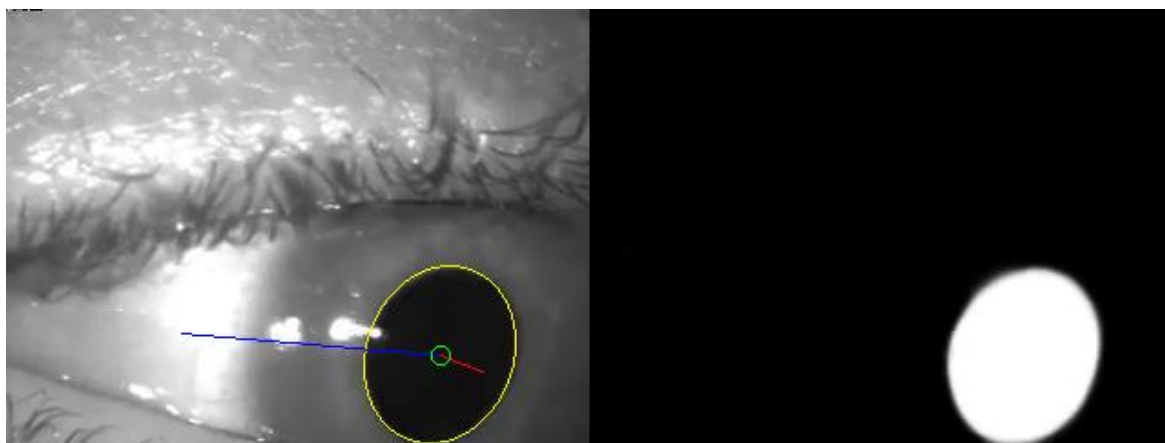
شکل ۴-۸ تصویر سوم خروجی الگوریتم (تصویر شماره دو).



شکل ۴-۹ تصویر چهارم خروجی الگوریتم (تصویر شماره سه).



شکل ۴-۱۰ تصویر پنجم خروجی الگوریتم (تصویر شماره چهار).



شکل ۴-۱۱ تصویر ششم خروجی الگوریتم (تصویر شماره پنج).

۴-۴- مشکلات و محدودیت‌ها

این الگوریتم تنها در شرایطی که تصاویر فقط شامل چشم فرد باشند عملکرد مناسبی دارد و در صورتی که تصویر شامل کل صورت باشد نمی‌تواند مردمک چشم را تشخیص دهد. در این الگوریتم فرض می‌شود دوربین در نقطه ثابتی نسبت به چشم قرار دارد، در نتیجه در طول آزمایش مجاز به حرکت دادن دوربین نیستیم. همچنین با توجه به لزوم ساخت مدل کره چشم قادر به تشخیص زاویه سمت نگاه بصورت بلادرنگ و یا فقط براساس یک عکس نمی‌باشد. اما در صورتی که مدل چشم فرد را از قبل داشته باشیم، می‌توانیم با سرعت نزدیک به بلادرنگ زاویه نگاه را بدست آوریم.

۴-۵- جمع‌بندی فصل

در این فصل ابتدا به شرح ساخت بخش سخت‌افزاری پروژه پرداختیم، سپس مشکلات و محدودیت‌های این سخت‌افزار را شرح دادیم. پس از آن ابزارها و بخش نرم‌افزاری پروژه را بررسی کردیم و در نهایت تعدادی از نتایج این پروژه را بررسی کردیم. به علت محدودیت‌های سخت‌افزاری امکان تست این نرم‌افزار در محیطی با کاربران واقعی وجود نداشت. به همین دلیل دقت پروژه محاسبه نشده است. هرچند از تصاویر تهیه شده به عنوان ورودی یکی از الگوریتم‌های ارائه شده پیشین [3] استفاده کردیم و مشاهده شد این الگوریتم قادر به تشخیص مردمک در شرایط نور محیطی عادی نمی‌باشد.

فصل پنجم

جمع‌بندی و پیشنهادات

جمع‌بندی و نتیجه‌گیری

در این پروژه می‌خواستیم برنامه‌ای جهت ردیابی جهت نگاه پیاده‌سازی کنیم. این برنامه دارای کاربردهای زیادی در حوزه‌های مختلف پزشکی، روان‌شناسی و مهندسی می‌باشد. در بخش نرم‌افزاری ما از ادغام دو الگوریتم برای بدست آوردن زاویه سمت نگاه چشم استفاده کرده‌ایم. این الگوریتم ابتدا در تصاویر تهیه شده از چشم فرد، با استفاده از روش ناحیه‌بندی مردمک چشم را تشخیص داده و جدا می‌کند. سپس یک مدل سه‌بعدی برای کره و مردمک چشم براساس دنباله‌ای از حرکات چشم بدست می‌آورد. این روش نیازی به واسنجی، شرایط خاص قرارگیری سر و چشم و نور پردازی تنظیم شده ندارد.

همچنین به کمک مقایسه با یکی از روش‌های پایه‌ای [3] دریافتیم که تصاویر دو بعدی مردمک که در به کمک الگوریتم ناحیه‌بندی بدست می‌آیند دارای دقت بهتری نسبت به تصاویر مردمک فعلی هستند اما توجه به این نکته الزامی است که تصاویر مردمک خوب الزاما باعث بهتر شدن عملکرد الگوریتم رهگیری سمت نگاه نمی‌شوند.

این الگوریتم در تشخیص تصاویر مردمک دقت خوبی دارد اما دقت آن به نسبت الگوریتم‌هایی که برای هر کاربر واسنجی دارند پایین‌تر است. [3] البته جبران این کاستی با جمع‌آوری اطلاعات بیشتر از چشم مانند واسنجی انجام می‌شود و بهبود و بهینه‌سازی الگوریتم تاثیر زیادی در بهبود دقت نخواهد داشت.

همچنین این الگوریتم در کد خود دقت را محاسبه نمی‌کند، و به علت محدودیت‌های سخت‌افزاری امکان محاسبه دقت در محیط آزمایشگاهی وجود نداشت. جهت آزمایش دقت این الگوریتم تصاویر تهیه شده توسط سخت‌افزار این پروژه را به عنوان ورودی یکی از نرم‌افزارهای پیشین استفاده کردیم. مشاهده کردیم که نرم‌افزار پیشین قادر به تشخیص مردمک چشم در این تصاویر نمی‌باشد. در نتیجه الگوریتم استفاده شده در این پروژه دارای دقت بیشتری در شرایط با نور تنظیم نشده می‌باشد.

پس از بدست آوردن نتایج ذکر شده می‌توانیم از این الگوریتم به عنوان فیلتر اول برای رهگیری سمت نگاه استفاده کنیم و با استفاده از جمع‌آوری داده‌های بیشتر و واسنجی به دقت بالاتری دست یابیم. در صورتی که برای هر کاربر واسنجی مجزا داشته باشیم می‌توانیم با دقت بیشتری مدل سازی سه‌بعدی

چشم را انجام دهیم. سپس از این مدل سه بعدی در الگوریتم‌های دیگر استفاده کنیم. برای مثال در مرحله ی بعد از مدل‌سازی از الگوریتم‌های شبکه عصبی برای تشخیص دقیق‌تر جهت نگاه استفاده کنیم.

همچنین می‌توانیم با توسعه این الگوریتم و ایجاد رابط کاربری مناسب از این پروژه برای معاینات پزشکی استفاده کنیم.

برای مثال این الگوریتم می‌تواند برای تشخیص میزان انحراف چشم مورد استفاده قرار گیرد. با توجه به دقت بالای الگوریتم استفاده از آن در معاینات پزشکی موجب کمتر شدن خطای انسانی در تشخیص می‌شود. چشم پزشک می‌تواند زاویه نگاه دو چشم بیمار در نگاه به نقاط مشخص شده را با هم مقایسه کرده و به میزان دقیق انحراف چشم پی ببرد. همچنین می‌توان در مراحل بعدی این پروژه الگوریتم دیگری پیاده‌سازی کرد که خود میزان انحراف چشم بیمار را محاسبه کرده و به کمک یک رابط کاربری مناسب به پزشک نمایش دهد.

از جمله کاربردهای دیگر این الگوریتم می‌توان به کنترل ویلچر برای افراد کم‌توان اشاره کرد. برای مثال می‌توان ویلچر را به گونه‌ای طراحی کرد که با تشخیص نگاه فرد بیش از چند ثانیه به جهتی خاص به سمت آن جهت حرکت کند و در صورتی که فرد به سمت پایین نگاه کرد توقف کند.

منابع و مراجع

- [1] S. Karthikeyan, V. Jagadeesh, R. Shenoy, M. Ecksteinz and B. Manjunath, "From Where and How to What We See," in *IEEE International Conference on Computer Vision*, Sydney, 2013.
- [2] A. Borji and L. Itti, "State-of-the-Art in Visual Attention Modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 185-207, 2013.
- [3] L. Swirski and N. Dodgson, "A fully-automatic, temporal approach to single camera," in *ECEM*, 2013.
- [4] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation.," *MICCAI*, vol. 3, pp. 234-241, 2015.
- [5] Y.-H. Yiu, M. Aboulatta, T. Raiser, L. Ophey, V. L. Flanagan, P. z. Eulenburg and S.-A. Ahmadi, "DeepVOG: Open-source pupil segmentation and gaze estimation in neuroscience using deep learning," *Journal of Neuroscience Methods*, vol. 324, 2019.
- [6] R. Safaee-Rad, I. Tchoukanov, K. C. Smith and B. Benhabib, "Three-dimensional location estimation of circular features for machine vision," in *IEEE Transactions on Robotics and Automation*, " *IEEE Transactions on Robotics and Automation*, vol. 8, no. 5, pp. 624-640, 1992.

Abstract

We present a 3D eye model fitting algorithm for use in gaze estimation , that operates on pupil ellipse geometry alone. It works with no user-calibration and does not require calibrated lighting. Our algorithm is based on fitting a consistent pupil motion model to a set of eye images . We describe a method of initializing this model from detected pupil ellipses , and a methods of iteratively optimizing the parameters of the model to best fit the original eye images.

Also, we have implemented a headset for eye-tracking. This headset is cheap and easy to use. And it can be used in future researches.

Key Words: gaze estimation, eye model, pupil detection, 3D eye-tracking



**Amirkabir University of Technology
(Tehran Polytechnic)**

Computer Engineering Department

B.Sc. Project

Implementation of 3D Eye tracking

**By
Yeganeh Kordi**

**Supervisor
Dr. Mohammad Rahmati**

October & 2021