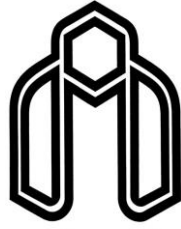


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی شاهرود
دانشکده برق و رباتیک

پایان نامه دوره کارشناسی ارشد مهندسی برق-الکترونیک

تعیین مکان چشم‌ها در تصاویر دو بعدی و ردیابی آنها

سعیده فردوسی

استاد راهنما

دکتر علیرضا احمدی فرد

استاد مشاور

دکتر علی سلیمانی ایوری

زمستان ۱۳۸۶

تشکر و قدردانی

اینجانب از جناب آقای دکتر احمدی فرد استاد راهنمای محترم، به خاطر زحمات و راهنمایی‌های بی‌دریغشان نهایت سپاسگزاری را دارم. بی شک نظرات و کمکهای ایشان تأثیر بسزایی در پیشبرد هر چه بهتر و سریعتر روند این پایان نامه داشته است.

فهرست مطالب

صفحه	عنوان
	فهرست جدول ها د
	فهرست شکل ها ه
	فصل ۱- مقدمه

 ۱
۱-۱-۱	معرفی سیستم های تشخیص وردیابی چشم ۱
۱-۲-۱	ساختمان و رفتار چشم انسان ۲
۱-۳-۱	تصویر برداری ۳
۱-۴-۱	کاربردها و موانع پیشروی الگوریتم های تشخیص و
۱-۵-۱	ردیابی چشم ۷
۱-۵-۱	سازماندهی پایان نامه ۹
	فصل ۲- مروری بر فعالیت های انجام شده ۱۰
۱-۲-۱	روش های تعیین مکان چشم ها در تصاویر دو بعدی .. ۱۱
۱-۱-۱-۲ روش های مبتنی بر
۱-۱-۱	ویژگی ۱۱
۱-۲-۱-۲ روش های مبتنی بر تطبیق
۱-۱-۱	الگو ۱۶
۱-۳-۱-۲ روش های مبتنی بر
۱-۲-۲	ظاهر ۲۰
۱-۲-۲	روش های ردیابی ۲۴
	فصل ۳- الگوریتم پیشنهادی برای تشخیص
	چشم ۴۶
۱-۳-۱	تعیین ساختار توپوگرافیک تصویر ۴۸
۱-۱-۱-۳ حذف نویز از تصویر
۱-۳-۱	ورودی ۴۹
۱-۲-۱-۳	استخراج برجسب های توپوگرافیک ۵۱
۱-۳-۱-۳	تعیین ساختار توپوگرافیک نمونه های آموزشی .. ۶۲
۱-۲-۳	استخراج ویژگی ها با استفاده از ممان های ثابت. ۶۸
۱-۲-۳	ممان ها وثابت های جبری ۷۰
۱-۲-۳ تابع مشخصه و تابع مولد
۱-۲-۳	ممان ۷۰
۱-۳-۲-۳ ممان
۱-۲-۳	مرکزی ۷۱
۱-۲-۳ محاسبه بردار
۱-۲-۳	ویژگی ها ۷۳
۱-۲-۳ تخمین پارامترهای مدل آماری برای کلاس های چشم
۱-۲-۳	راست، چشم چپ و نواحی غیر چشم ۷۴
۱-۳-۳	تشخیص محل چشم در تصویر آزمون ۷۵
۱-۳-۳ تعیین ساختار توپوگرافیک تصویر
۱-۲-۳	ورودی ۷۵

۳-۳-۲-	تعیین نقاط کاندید ای حضور مردمک	۷۶
۳-۳-۳-	گروه بندی کاندیدها و حذف تعدادی از آنها	۸۱
۳-۳-۴-	تعیین محل مردمک و جداسازی ناحیه چشم از تصویر ورودی	۸۵
۴-	فصل ۴- رديابي چشم در يك دنباله ويديويي	۹۴
۴-۱-	مروری بر روش‌های رديابي	۹۴
۴-۱-۱-	رديابي با استفاده از الگو و مدل‌های مبتنی بر چگالی	۹۵
۴-۱-۲-	الگوریتم تطبیق بلوک (BMA)	۱۰۰
۴-۲-	الگوریتم پیشنهادی برای رديابي چشم‌ها در دنباله تصاویر	۱۰۱
۵-	فصل ۵- تعیین فاصله بین پلك‌ها در چشم	۱۰۴
۵-۱-	توابع نگاشت	۱۰۶
۵-۱-۱-	تابع نگاشت جمع کننده	۱۰۶
۵-۱-۲-	تابع نگاشت واریانس	۱۰۷
۵-۱-۳-	تشخیص نواحی مختلف چشم توسط توابع نگاشت	۱۰۸
۵-۲-	الگوریتم پیشنهادی برای تعیین فاصله بین پلك‌ها	۱۱۲
۶-	فصل ۶- نتایج تجربی	۱۲۰
۶-۱-	پایگاه داده	۱۲۰
۶-۱-۱-	داده‌های آموزشی	۱۲۰
۶-۱-۲-	پایگاه داده مربوط به تست الگوریتم	۱۲۲
۶-۲-	نتایج	۱۲۳
۶-۲-۱-	نتایج الگوریتم تشخیص چشم	۱۲۳
۶-۲-۲-	نتایج الگوریتم رديابي چشم	۱۳۴
۶-۲-۳-	نتایج الگوریتم تعیین فاصله بین پلك‌ها	۱۴۶
	نتیجه گیری و پیشنهادات	۱۵۳
	ضمیمه الف	۱۵۶

فهرست جدول‌ها

صفحه

عنوان

- جدول ۱-۲ : نتایج بدست آمده در روش Zhou و همکارش [۸]..... ۱۴
- جدول ۲-۲ : نتایج بدست آمده در روش Black و همکارانش [۲۵]..... ۲۶
- جدول ۳-۲ : نتایج بدست آمده در روش Haro و همکارانش [۳۰]..... ۳۲
- جدول ۱-۳ : تعیین برجسب‌های توپوگرافیک تصویر با استفاده از روش Wang و همکارش [۴۵]..... ۶۱
- جدول ۱-۶ : مقایسه نرخ آشکارسازی برای روش Wang و همکارش [۴۷] و الگوریتم پیشنهادی ۱۲۸
- جدول ۲-۶ : مقایسه سرعت برای روش Wang و همکارش [۴۷] و الگوریتم پیشنهادی ۱۲۸
- جدول ۳-۶ : نتایج الگوریتم پیشنهادی..... ۱۳۱

فهرست شکل‌ها

صفحه	عنوان
۲	شکل ۱-۱: ویژگی‌های چشم.....
۴	شکل ۲-۱: نتایج تصویر برداری با نور مادون قرمز.....
۵	شکل ۳-۱: اثر انعکاس قرنیه‌ای در چشم.....
۶	شکل ۴-۱: سیستم تشخیص جهت نگاه.....
۶	شکل ۵-۱: سیستم تصویر برداری با نور مادون قرمز.....
۱۷	شکل ۱-۲: الگوی چشم در روش Orazio و همکارانش [۱۳].....
۱۸	شکل ۲-۲: الگوی چشم در روش Tan و همکارانش [۱۴].....
۳۸	شکل ۳-۲: تعیین نقاط کاندید مردمک در تصویر باینری.....
۴۱	شکل ۴-۲: شبکه عصبی مورد استفاده در روش Marcne و همکارانش [۳۸].....
۴۲	شکل ۵-۲: نیم دایره استفاده شده به عنوان ماسک.....
۴۳	شکل ۶-۲: نتایج تشخیص و ردیابی عنبیه در روش پیشنهادی Tan و همکارش [۴۲].....
۴۳	شکل ۷-۲: نتایج تشخیص و ردیابی عنبیه در روش Tian و همکارش [۴۰].....
۴۴	شکل ۸-۲: نتایج ردیابی پلک‌ها در روش Tian و همکارش [۴۰].....
۴۵	شکل ۹-۲: نتایج ردیابی پلک‌ها در روش Tan و همکارش [۴۲].....
۴۷	شکل ۱-۳: مراحل مختلف الگوریتم پیشنهادی.....
۴۸	شکل ۲-۳: نمایش تصویر چهره بصورت یک رویه سه بعدی.....
۴۹	شکل ۳-۳: رویه متناظر با چشم راست.....
۴۹	شکل ۴-۳: الف) تصویر اصلی چشم و رویه مربوط به آن (ب) تصویر چشم پس از اعمال فیلتر گوسی و رویه مربوط به آن.....
۵۳	شکل ۵-۳: فیلترهای گسسته دو بعدی برای $N=20$
۵۵	شکل ۶-۳: برچسب‌های توپوگرافیک.....
۵۹	شکل ۷-۳: وقوع ماکزیمم محلی در مکانی بجز مرکز پیکسل.....
۵۹	شکل ۸-۳: وقوع مینیمم محلی در مکانی بجز مرکز پیکسل.....
۶۰	شکل ۹-۳: تصحیح بردارهای ویژه.....
۶۲	شکل ۱۰-۳: تصاویر مورد استفاده در آموزش سیستم.....
۶۳	شکل ۱۱-۳: الف) تصویر اولیه از چشم راست (ب) تصویر نگاشت یافته به حوزه توپوگرافیک (ج) تفکیک برچسب‌های ظاهر شده در تصویر نگاشت یافته.....
۶۳	شکل ۱۲-۳: الف) تصویر اولیه از چشم چپ (ب) تصویر نگاشت یافته به حوزه توپوگرافیک (ج) تفکیک برچسب‌های ظاهر شده در تصویر نگاشت یافته.....
۶۴	شکل ۱۳-۳: الف) تصویر اولیه غیر چشم (ب) تصویر نگاشت یافته به حوزه توپوگرافیک (ج) تفکیک برچسب‌های ظاهر شده در تصویر نگاشت یافته.....

- شکل ۳-۱۴: الف) تصویر اولیه غیر چشم ب) تصویر نگاشت یافته به حوزه توپوگرافیک ج) تفکیک برچسب‌های ظاهر شده در تصویر نگاشت یافته ۶۵
- شکل ۳-۱۵: الف) تصویر چشم راست ب) فراوانی برچسب‌ها در تصویر برچسب خورده توپوگرافیکی ۶۶
- شکل ۳-۱۶: الف) تصویر چشم چپ ب) فراوانی برچسب‌ها در تصویر برچسب خورده توپوگرافیکی ۶۶
- شکل ۳-۱۷: الف) تصویر غیرچشم ب) فراوانی برچسب‌ها در تصویر برچسب خورده توپوگرافیکی ۶۷
- شکل ۳-۱۸: الف) تصویر غیرچشم ب) فراوانی برچسب‌ها در تصویر برچسب خورده توپوگرافیکی ۶۷
- شکل ۳-۱۹: تعیین ساختار توپوگرافیک تصویر آزمون ۷۶
- شکل ۳-۲۰: ویژگی توپوگرافیکی مردمک ۷۷
- شکل ۳-۲۱: بررسی نقطه کاندید مردمک در شرایط مختلف ۷۹
- شکل ۳-۲۲: تعیین نقاط کاندیدای مردمک ۸۰
- شکل ۳-۲۳: استفاده از برچسب سرایشی بدون انحناء در دسته‌بندی و حذف کاندیدهای نادرست ۸۲
- شکل ۳-۲۴: خصوصیات هندسی مربوط به نقاط کاندید ۸۳
- شکل ۳-۲۵: توزیع برچسب ناودانی بیرون در تصویر برچسب خورده توپوگرافیکی شکل ۳-۲۳ الف) ... ۸۵
- شکل ۳-۲۶: مراحل کلاسه‌بندی مناطق کاندید چشم در تصویر ۸۸
- شکل ۳-۲۷: نتایج الگوریتم پیشنهادی در شرایط مختلف ۹۰
- شکل ۴-۱: ردیابی در روش Adachi و همکارش [۵۷] ۹۷
- شکل ۴-۲: تصویرهای مربوط به تکرارها در روش جابجایی میانگین ۹۹
- شکل ۴-۳: بلوک دیاگرام الگوریتم پیشنهادی برای ردیابی چشم‌ها ۱۰۳
- شکل ۵-۱: تصاویر تغییر حالت چشم در پلک زدن ۱۰۴
- شکل ۵-۲: تصویر چشم در سطح خاکستری ۱۰۵
- شکل ۵-۳: عملکرد مناسب VPF_v در تعیین مرزهای شکل ۱۰۷
- شکل ۵-۴: عملکرد مناسب IPF_v در تعیین مرزهای شکل ۱۰۸
- شکل ۵-۵: مدل چشم ۱۰۸
- شکل ۵-۶: تعیین مرزهای چشم توسط تابع نگاشت جمع کننده ۱۰۹
- شکل ۵-۷: تعیین مرزهای چشم ۱۱۱
- شکل ۵-۸: تشخیص باز یا بسته بودن چشم‌ها ۱۱۲
- شکل ۵-۹: شناسایی مرزهای افقی چشم ۱۱۴
- شکل ۵-۱۰: پاسخ تابع نگاشت عمودی در مکان پلک‌ها ۱۱۵
- شکل ۵-۱۱: تشخیص مکان پلک‌ها ۱۱۶
- شکل ۵-۱۲: نتایج الگوریتم ردیابی چشم در فرایند پلک زدن ۱۱۷
- شکل ۵-۱۳: تصاویر اندازه‌گردان روشنایی در جهت عمودی ۱۱۸
- شکل ۵-۱۴: منحنی نمایش دهنه فاصله پلک‌ها در پلک زدن ۱۱۸
- شکل ۵-۱۵: منحنی تابع نگاشت جمع کننده در جهت افقی برای شکل شماره ۴ ۱۱۹
- شکل ۵-۱۶: منحنی تابع نگاشت جمع کننده در جهت افقی برای شکل شماره ۱۰ ۱۱۹

- شکل ۵-۱۷ : منحنی تابع نگاشت جمع کننده در جهت افقی برای شکل شماره ۱۳..... ۱۱۹
- شکل ۶-۱ : پایداری چاله‌های مربوط به مردمک در اثر تغییر پارامترهای فیلتر هموارساز گوسی..... ۱۲۴
- شکل ۶-۲ : مقایسه نتایج در اثر تغییر ابعاد پنجره فیلتر مشتق‌گیر..... ۱۲۵
- شکل ۶-۳ : مقایسه نتایج در اثر تغییر پارامترهای فیلتر هموارساز گوسی..... ۱۲۶
- شکل ۶-۴ : نتایج حاصل از پیاده‌سازی الگوریتم Wang و همکارش [۴۷]..... ۱۲۹
- شکل ۶-۵ : نتایج الگوریتم پیشنهادی..... ۱۳۱
- شکل ۶-۶ : تعدادی از نتایج بدست آمده توسط الگوریتم پیشنهادی..... ۱۳۳
- شکل ۶-۷ : تعدادی از نتایج خطا در الگوریتم پیشنهادی..... ۱۳۴
- شکل ۶-۸ : نتایج الگوریتم ردیابی چشم در حرکت انتقالی سر..... ۱۳۶
- شکل ۶-۹ : نتایج الگوریتم ردیابی چشم در چرخش عمقی سر..... ۱۳۸
- شکل ۶-۱۰ : نتایج الگوریتم ردیابی در حرکت انتقالی و چرخش عمقی سر..... ۱۴۰
- شکل ۶-۱۱ : نتایج الگوریتم ردیابی در حرکت انتقالی و چرخش عمقی سر..... ۱۴۱
- شکل ۶-۱۲ : نتایج الگوریتم ردیابی در استفاده از عینک..... ۱۴۳
- شکل ۶-۱۳ : نتایج الگوریتم ردیابی در پلک زدن..... ۱۴۴
- شکل ۶-۱۴ : نتایج الگوریتم ردیابی در حرکت سطحی سر در صفحه دوربین..... ۱۴۶
- شکل ۶-۱۵ : تعیین مکان پلک‌ها در پلک زدن..... ۱۴۷
- شکل ۶-۱۶ : منحنی نمایش دهنه فاصله پلک‌ها در پلک زدن..... ۱۴۸
- شکل ۶-۱۷ : تعیین فاصله بین پلک‌ها در خواب‌آلودگی..... ۱۴۸
- شکل ۶-۱۸ : منحنی نمایش دهنه فاصله پلک‌ها در خواب‌آلودگی..... ۱۴۹
- شکل ۶-۱۹ : پاسخ الگوریتم در رابطه با تصاویر همراه با عینک..... ۱۵۰
- شکل ۶-۲۰ : پاسخ الگوریتم پس از تصحیح ناحیه چشم..... ۱۵۰
- شکل ۶-۲۱ : پاسخ الگوریتم در رابطه با انعکاس نور در شیشه عینک..... ۱۵۱
- شکل ۶-۲۲ : پاسخ الگوریتم در رابطه با انعکاس نور در شیشه عینک..... ۱۵۲

چکیده

تشخیص و ردیابی چشم‌ها نقش کلیدی در توسعه سیستم‌های مربوط به تعامل بین انسان و کامپیوتر (HCI) دارد. سیستم‌های یاری رسان به افراد معلول، تشخیص خواب آلودگی راننده، تشخیص هویت افراد از جمله سیستم‌های مذکور می‌باشند. کارائی این سیستمها در گرو تشخیص درست محل چشم‌ها و ردیابی آنها در یک رشته ویدئویی می‌باشد. مهمترین مشکلاتی که عملکرد الگوریتم‌های تشخیص و ردیابی چشم‌ها را دچار چالش می‌کند عبارتند از چرخش سر در عمق و در صفحه عمود بر محور اپتیکی دوربین، پوشیده شدن چشم‌ها توسط عینک و مو و شرایط بد نورپردازی می‌باشد. تحقیقات زیادی برای حل مشکلات ذکر شده در حال انجام است.

در این پایان‌نامه بدنبال ارائه روشی جهت تشخیص محل چشمها و سپس ردیابی آنها در یک رشته ویدئویی هستیم. در این روش بر مقاوم سازی سیستم پیشنهادی، در برابر تبدیل هندسی تشابهی و شرایط بد نورپردازی تأکید شده است. برای این منظور بجای استفاده مستقیم از روشنایی تصویر در ناحیه چشم از ساختار توپوگرافیکی تصویر استفاده می‌کنیم. تصویر برچسب خورده توپوگرافیکی را به کمک توابع تغییر ناپذیر با تبدیل هندسی تشابهی توصیف می‌کنیم. این توابع بر حسب ممان‌های نواحی برچسب خورده تعریف می‌گردند. به کمک تصاویر چشم راست، چشم چپ و غیر چشم، سیستم تشخیص چشم‌ها را آموزش می‌دهیم. در بخش دوم تحقیق، روشی جهت ردیابی چشم‌ها در رشته ویدئویی پیشنهاد می‌دهیم. الگوریتم با حداکثر نمودن معیار همبستگی متقابل بین الگوی چشم‌ها در فریم‌های اولیه و فریم‌های بعدی، چشم‌ها را ردیابی می‌کند. الگوریتم پیشنهادی از پیچیدگی کمی برخوردار می‌باشد. در این تحقیق همچنین الگوریتمی جهت تعیین فاصله بین پلک‌ها ارائه شد. در این روش از نگاشت گرادیان تصویر چشم در راستای افقی استفاده شد.

نتایج الگوریتم تشخیص چشمها بر روی پایگاه داده XM2VTS کارائی خوب روش را در شرایط مختلف تصویر برداری نشان می‌دهد. همچنین نتیجه آزمایش روش ردیابی بر روی رشته‌های ویدئویی نشان می‌دهد که چنانچه سر چرخش عمقی قابل توجهی نداشته باشد ردیابی با موفقیت انجام می‌شود.

فصل ۱ - مقدمه

در این تحقیق مسأله تشخیص چشم و ردیابی آن در تصاویر دوبعدی مورد بررسی قرار می‌گیرد. با توجه به گسترش روزافزون کاربردهای تشخیص و ردیابی چشم و مشکلاتی که بر سر راه این الگوریتم‌ها وجود دارد، نیاز به روش‌های دقیق و تا حد امکان مقاوم کاملاً احساس می‌شود. در این تحقیق تلاش شده‌است با ارائه روش‌های نو و کارآمد گامی در جهت بهبود عملکرد چنین الگوریتم‌هایی برداشته شود. در قسمت آخر نیز نگاهی به مطالعه تعیین فاصله بین پلک‌ها که اساسی‌ترین قدم در پاسخ به مسائلی از قبیل تشخیص خواب‌آلودگی یا تشخیص پلک‌زدن می‌باشد، خواهیم داشت.

۱-۱ - معرفی سیستم‌های تشخیص و ردیابی چشم

هدف الگوریتم‌های تشخیص چشم در تصاویر، پاسخ به این سؤال است که آیا چشم در تصویر وجود دارد یا نه؟ و در صورت مثبت بودن پاسخ، مکان و محدوده هر چشم تعیین شود. چشم دارای ویژگی‌های منحصر بفردی است که جستجوی این ویژگی‌ها در تصویر، یافتن پاسخ به این سؤال را امکان‌پذیر می‌نماید. منظور از ردیابی چشم در یک دنباله ویدئویی^۱ دنبال کردن مکان چشم در فریم‌های متوالی از آن دنباله می‌باشد.

استفاده از الگوریتم‌های بینایی ماشین و پردازش تصویر در کنار سخت‌افزار مناسب امکان استخراج ویژگی‌های لازم برای تشخیص چشم و ردیابی آن را در فریم‌های متوالی فراهم می‌سازد. روند انجام چنین پروسه‌ای بصورت زیر می‌باشد.

یک دوربین CCD بطور زنده از صحنه تصویر برداری می‌کند و رشته تصاویر بصورت دیجیتالی ضبط و به کامپیوتر ارسال می‌گردند. این کامپیوتر می‌تواند با قابلیت کامپیوترهای شخصی معمولی و یا مجهز به پردازنده DSP جهت افزایش سرعت اجرای الگوریتم‌ها باشد. الگوریتم‌های بینایی ماشین و پردازش تصویر بر روی کامپیوتر اجرا شده، مکان چشم را در تصویر مربوط به فریم اول در رشته ویدئویی تشخیص داده و با ورود فریم‌های بعدی مکان چشم را ردیابی می‌کنند.

با استفاده از پردازش تصویر یک کامپیوتر به تنهایی و بدون نیاز به کمک انسان، تبدیل به ابزاری مفید برای جمع‌آوری و تحلیل اطلاعات دیداری (بصری)^۲ می‌شود. از نظر تئوری یک سیستم بینایی ماشین پتانسیل عملکرد برابر و در بعضی موارد حتی با قدرت بیشتر از بینایی انسان را دارد با این تفاوت که در آن اثری از کسالت و خستگی وجود ندارد. البته کاملاً واضح است که رسیدن به چنین سیستمی نیازمند مطالعات و تحقیقات بسیار است.

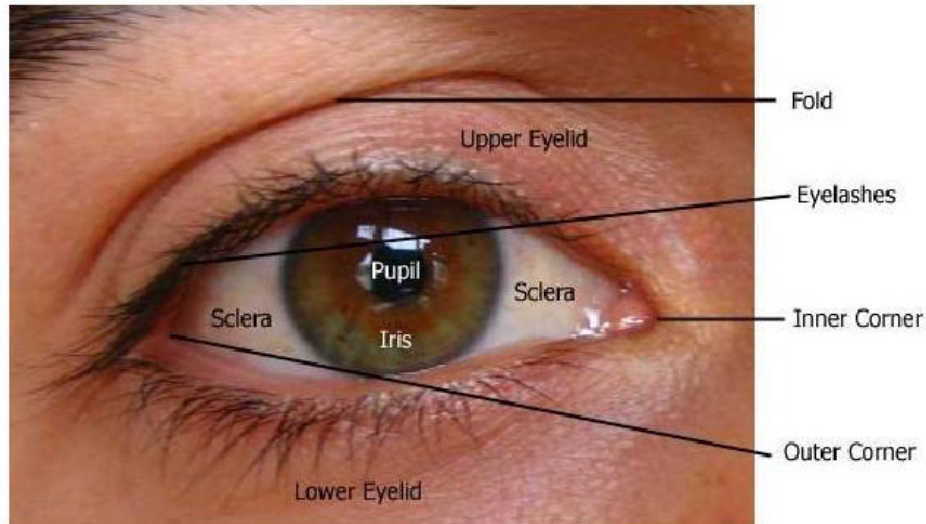
همانطور که قبلاً نیز گفته شد تشخیص چشم مستلزم تشخیص ویژگی‌های چشم می‌باشد. قسمت بعد به بیان اطلاعات جامعی در رابطه با ویژگی‌های چشم و حرکت آنها در ناحیه چشم می‌پردازد.

¹ Video Clip

² Visual

۲-۱- ساختمان و رفتار چشم انسان

شکل ۱-۱ نمای نزدیک^۱ یک چشم باز را نشان می‌دهد. یکی از مهمترین اجزاء چشم عنبیه^۲ می‌باشد. عنبیه دارای یک ساختار دایره‌ای شکل است و رنگ و بافت آن در محدوده وسیعی تغییر می‌کند.



شکل ۱-۱: ویژگی‌های چشم

حتی اگر چشم در وضعیت نرمال (چشم باز یا چشمی که با عوامل خارجی پوشانده نشده است) تصویر برداری شود، امکان دارد دایره محیطی عنبیه بطور کامل در تصویر دیده نشود. قابل رؤیت بودن عنبیه تحت تأثیر عوامل مختلفی می‌باشد. اکثر مواقع بخشی از عنبیه توسط پلک بالا یا پلک پایین یا هر دو پلک پوشانده می‌شود. مکان این ویژگی نیز تغییر می‌کند و می‌تواند به یکی از دو گوشه چشم یا بسمت بالا و پایین منتقل شود. سرعت حرکت عنبیه نیز معرف ویژگی دیگری از چشم می‌باشد که بیانگر اطلاعات مفیدی در رابطه با نوع رفتار شخص است.

مردمک^۳ یکی دیگر از اجزای چشم است که داخل عنبیه قرار دارد و بصورت یک دایره کوچک تیره در تصویر دیده می‌شود. اندازه مردمک وابسته به شرایط نوری محیط می‌باشد.

قسمتی از کره چشم که با رنگ سفید دیده می‌شود، صلبیه^۴ نام دارد که بالاترین سطح روشنایی را در ناحیه چشم بخود اختصاص داده است. این جزء عنبیه را بطور کامل احاطه می‌کند. در کنار اجزاء مذکور، پلک بالا و پلک پایین دو جزء مهم دیگر هستند که در تشخیص چشم‌ها در تصاویر مؤثر می‌باشند.

در حالت کلی حرکت اجزای چشم نسبت به هم وضعیت پیچیده‌ای دارد که می‌تواند نمایانگر بخشی از رفتار و عکس‌العمل فرد به محیط اطراف باشد.

¹ Close up

² Iris

³ Pupil

⁴ Sclera

۳-۱- تصویر برداری

تصاویر تهیه شده توسط سنسور بینایی که می‌تواند یک دوربین ویدئویی باشد داده‌های ورودی به یک سیستم پردازش تصویر را فراهم می‌سازد. نوع داده‌های (تصاویر) ورودی در سیستم‌های تشخیص و ردیابی چشم به دو دسته مهم تقسیم می‌شوند:

۱. تصاویر گرفته شده توسط سنسورها و منابع مادون قرمز

۲. تصاویر گرفته شده با استفاده از سنسورهای نور مرئی

تحقیق انجام شده در این پایان‌نامه بر اساس تصویر برداری با نور مرئی می‌باشد، یک از مزایای مهم این روش عدم نیاز به تجهیزات نوری مادون قرمز و استفاده از دوربین‌های معمولی می‌باشد. برای آشنایی، در این قسمت روش تصویر برداری نور مادون قرمز مختصراً توضیح داده می‌شود. برای افزایش کنتراست بین مردمک و عنبیه برخی از سیستم‌های تشخیص و ردیابی چشم‌ها از منبع نوری و دوربین مادون قرمز استفاده می‌کنند. این نور قابل رویت نیست و برای کاربر مزاحمت ایجاد نمی‌کند. در عمل اکثر این سیستم‌ها از امواج مادون قرمز با طول موج حدود 880 nm که غیر قابل رویت توسط کاربر ولی قابل تشخیص توسط اکثر دوربین‌های معمولی می‌باشد استفاده می‌کنند. در برخی مواقع منبع نور مادون قرمز^۱ IR نزدیکی محور نوری دوربین قرار می‌گیرد. به دلیل اینکه در این حالت دوربین قادر به دریافت نور مادون قرمز منعکس شده از پشت چشم می‌باشد، مردمک در تصویر روشن خواهد بود. اما اگر منبع نور مادون قرمز دور از محور نوری دوربین واقع شود یا تصویر برداری با نور مرئی انجام شود مردمک در تصویر تیره خواهد بود (شکل ۱-۲).

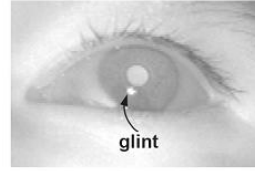
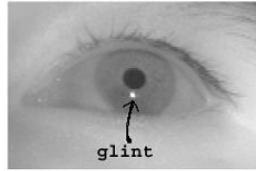


شکل ۱-۲: نتایج تصویر برداری با نور مادون قرمز

همچنین وجود منبع نور مادون قرمز^۲ یک انعکاس قرنی^۲ یا درخشش در قرنی^۲ را در تصویر تولید می‌کند که به خوبی در شکل ۱-۳ قابل رویت است. از این نقطه نیز می‌توان برای پی بردن به محل مردمک استفاده کرد.

^۱ Infra red

^۲ Corneal Reflection CR



(ب): تصویر مردمک تیره

(الف): تصویر مردمک روشن

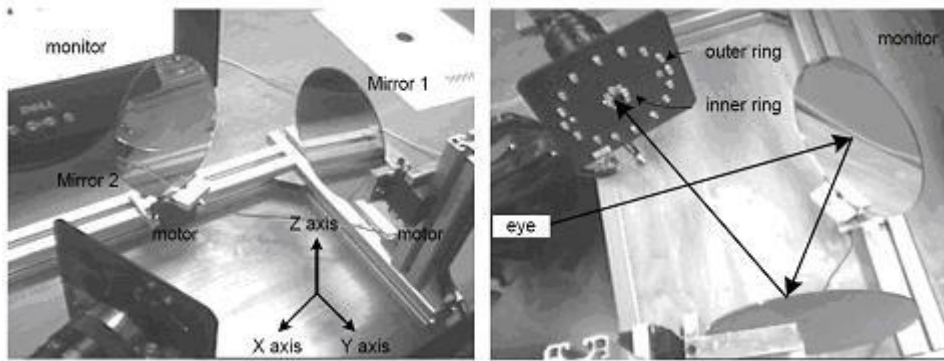
شکل ۱-۳: اثر انعکاس قرنیهای در چشم

Reulen و همکارانش [۱] با استفاده از نور مادون قرمز به ردیابی مرز بین عنبیه و صلبیه^۱ پرداخته‌اند. بدلیل اختلاف زیادی که بین کنتراست دو ناحیه صلبیه و عنبیه وجود دارد مرز بین آن دو بسادگی در حالت افقی ردیابی می‌شود اما با پوشانده شدن این مرز توسط پلک‌ها در بالا و پایین چشم، ردیابی آن در حالت عمودی نتیجه قابل قبولی ندارد. در این روش از دی‌ودهای مادون قرمز و ترانزیستورهای نوری به ترتیب در بالا و پایینی چشم استفاده می‌کنند. چندی‌ن جفت از این ترکیب به گونه‌ای نصب می‌گردد که مرز بین عنبیه و صلبیه همواره از جهات مختلف بینی و گیجگاه نوردهی یکسانی شود. ترانزیستورهای نوری، نور IR منعکس شده را به ولتاژ تبدیلی می‌کنند. ولتاژ تولید شده مربوط به نور منعکس شده در جهت بینی با ولتاژ مربوط به انعکاس در جهت گیجگاه مقایسه می‌شود، اختلاف این دو ولتاژ زاویه انحراف چشم را نشان می‌دهد. دقت بالا و پهنای باند وسیع در این روش امکان مطالعات اساسی رفتار چشم را برای متخصصان فراهم آورده است. این سیستم بروی سر انسان سوار شده و وزنی حدود ۳۰۰ گرم دارد و عموماً برای معالجات بالینی چشمها در انسان استفاده می‌گردد.

نوع دیگری از سیستمهای تشخیصی ص دهندۀ جهت چشم (EGT)^۲ با هدف حذف حرکات سر کاربر و کاهش محدودیت بی حرکت بودن کاربر توسط Kim و همکارانش ارائه شد [۲]. این سیستم متشکل از یک دوربین مادون قرمز و دو آینه با قطر ۱۵ سانتی‌متر می‌باشد (شکل ۱-۴) که در مقابل فرد قرار می‌گیرد. دوربین تصاویر چشم منعکس شده از آینه ۱ و ۲ را ثبت می‌کند. آینه ۱ حول محور Y می‌چرخد و جبران کننده حرکات سر در راستای عمودی است. آینه ۲ حول محور Z می‌چرخد و برای جبران سازی حرکات سر در راستای افقی قرار داده شده است. زوایای عنوان شده با استفاده از دو موتور پله‌ای که به کامپیوتر متصل هستند قابل کنترل می‌باشند. استفاده از این آینه‌ها که برای جبران حرکات سر تعبیه شده اند اختیاری است و در برخی از سیستمها استفاده می‌گردد.

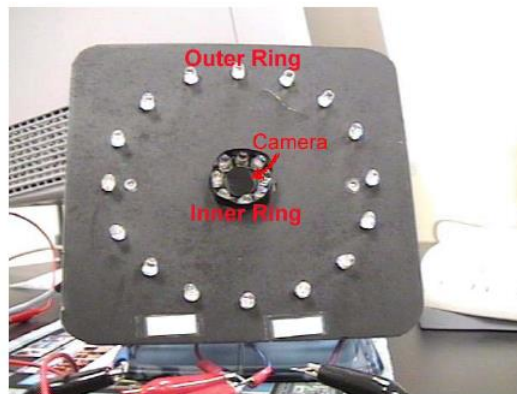
¹ Limbus

² Eye Gaze Tracker



شکل ۱-۴: سیستم تشخیص جهت نگاه

در این سیستم از منبع نور مادون قرمز برای تعیین موقعیت مردمک استفاده می‌گردد. دو حلقه از LED های مادون قرمز که به عنوان منبع نور، ماکزیمم طول موج آنها حدود 800nm است در اطراف دوربین تعبیه شده‌اند (شکل ۱-۵). برای حداقل ساختن تداخل سایر منابع نوری موجود در محیط از فیلمتر پایینی گذر نوری نیز استفاده می‌گردد. هنگامی که LED های حلقه داخلی روشن می‌شوند، دوربین تصویری مردمک روشن را می‌گیرد و هنگامی که LED های حلقه خارجی روشن می‌شوند تصویری مردمک بصورت لکه تیره حاصل می‌گردد. روشن و خاموش شدن حلقه داخلی و خارجی از روی سیگنالهای جداساز می‌دان فرد و زوج که از دوربین CCD حاصل می‌شود کنترل می‌گردند.



شکل ۱-۵: سیستم تصویر برداری با نور مادون قرمز

۴-۱- کاربردها و موانع پیشروی الگوریتم‌های تشخیص و ردیابی چشم

تهیه داده‌های مناسب که بیانگر ویژگی‌های مؤثر در تصمیم‌گیری الگوریتم‌های بینایی ماشین هستند، دستیابی به اهداف ذیل را امکان‌پذیر می‌سازند:

- سیستم‌های تشخیص و شناسایی چهره^۱: چشمها نمایان‌ترین و پایدارترین ویژگی را در

چهره افراد دارند، بنابراین تشخیص چشمها اغلب بعنوان مهمترین قدم در بسیاری از

الگوریتم‌های تشخیص چهره محسوب می‌شود. همچنین بسیاری از سیستم‌های تشخیص

¹ Face detection and Face recognition

چهره بر اساس شناسایی اجزای اصلی چهره مثل چشم، بینی، دهان و ارتباط هندسی آنها طراحی می‌شوند [۳]. در سیستم‌های تشخیص چهره، ابتدا باید یکنواخت سازی تصاویر از نظر اندازه و جهت چهره صورت پذیرد، همچنین در بسیاری از سیستم‌ها، شناسایی چهره بر اساس شناسایی چشم‌ها در تصویر انجام می‌گیرد چون چشم‌ها اجزاء متمایز در چهره می‌باشند که نسبت فاصله آنها به ابعاد چهره تقریباً ثابت است.

- **سیستم‌های دنبال کننده جهت نگاه افراد^۱**
 - **روش‌های مربوط به تعامل بین انسان و کامپیوتر^۲ (HCI) :** در حضور فناوریهای نو و رسانه‌ها نیاز به توسعه روشهای مؤثر به منظور تعامل بین انسان و کامپیوتر (HCI) کاملاً احساس می‌شود. چنین روش‌هایی مستقل از روشهای ارتباطی قدیمی که توسط وسایلی از قبیل صفحه کلید، میکروفن و صفحه‌نمایش عمل می‌کنند، می‌باشند. بدلیل اینکه چشمها در بیشتر موارد میل و خواسته فرد را منعکس می‌کنند، تشخیص و ردیابی آنها به منظور تحلیل حالات چهره و واکنش بین انسان و کامپیوتر تعیین کننده است [۴].
 - **سیستم‌های ناظر بر هشیاری:** یکی از موارد کاربرد تشخیص چشمها همراه با ردیابی آنها، تشخیص خواب آلودگی افراد است. تعیین خواب آلودگی و میزان آن اهمیت بسیار زیادی در مطالعه خواب و تشخیص بیماریها دارد. همچنین از تشخیص خواب آلودگی می‌توان برای هشدار رانندگان و جلوگیری از حوادث رانندگی استفاده نمود [۵].
 - **سیستم‌های یاری رسان به افراد معلول:** در چنین سیستم‌هایی پیاده‌سازی الگوریتم بنحوی انجام شده‌است که عکس‌العمل مناسبی در مقابل رفتار شخص معلول نشان خواهد داد.
- چالش‌ها و مشکلات همراه با تشخیص و ردیابی چشم به موارد مشخص شده در زیر منسوب می‌شوند:
- **وضعیت سر^۳ :** با توجه به وضعیت قرار گرفتن نسبی سر و دوربین که باعث تغییر زاویه تصویر برداری می‌شود حالت‌های مختلفی از قبیل نیمرخ، تمامرخ، وارونگی چهره در تصویر و چرخش

¹ Gaze detection and tracking

² Human Computer Interaction (HCI)

³ Pose

۴۵ درجه سر در تصویر حاصله ایجاد می‌شود. در چنین وضعیتی ممکن است اجزاء چشم بخوبی قابل تشخیص نباشند.

- پوشیده شدن بخشی از چشم‌ها : حضور اجزایی از قبیل عینک که رنگ و اندازه آن نیز تغییر می‌کند، می‌تواند عملکرد سیستم‌های تشخیص و ردیابی چشم را تحت تأثیر قرار دهد.
- حالات چهره^۱: نمود و ظاهر اجزای چهره در تصاویر مستقیماً متأثر از حالات مختلف صورت افراد است، مثلاً به هنگام اخم و عصبانیت حالت چشم‌ها تغییر می‌کند.
- چرخش در عمق : با چرخش سر نسبت به محور نوری دوربین، وضعیت چشم‌ها در تصویر حاصله مستقیماً تغییر می‌کند و در صورتیکه الگوریتم نسبت به چرخش مقاوم نباشد، نتیجه تشخیص و ردیابی همراه با خطاست.
- شرایط عکسبرداری: زمانی که یک تصویر ساخته می‌شود، فاکتورهایی از قبیل نورپردازی (طیف نور، توزیع منبع و شدت نور) و ویژگی‌های دوربین (پاسخ سنسور و لنز) نحوه ظهور چشم را در تصاویر تحت تأثیر قرار می‌دهند.

۱-۵- سازماندهی پایان‌نامه

در فصل دوم مروری بر فعالیت‌ها و تحقیقات گزارش شده در زمینه تشخیص و ردگیری چشم‌ها ارائه می‌گردد. مطالب بیان شده در این فصل شامل خلاصه‌ای از چگونگی عملکرد روش‌های پیشنهاد شده، نقاط ضعف و قوت آنها و نتایج بدست آمده توسط این روش‌ها می‌باشد.

در فصل سوم روش پیشنهادی برای تشخیص چشم ارائه خواهد شد. در طراحی این روش سعی شده است تا حد امکان مشکلات موجود در رابطه با تشخیص چشم برطرف گردد و با افزایش دقت و مقاومت در الگوریتم پیشنهادی بهبود مناسبی در عملکرد سیستم و نتایج حاصله ایجاد شود.

در فصل چهارم الگوریتم پیشنهادی در رابطه با ردیابی چشم‌ها در دنباله ویدئویی شرح داده می‌شود.

در فصل پنجم الگوریتم پیشنهاد شده برای تعیین فاصله پلک‌ها ارائه می‌شود. در این فصل روشی برای محاسبه فاصله پلک‌ها نسبت به زمان در طول یک رشته ویدئویی ارائه می‌گردد. این مطالعه می‌تواند در تشخیص میزان هوشیاری راننده و مطالعات مربوط به خواب مفید باشد.

¹ Facial expression

در فصل ششم نتایج حاصل از اعمال الگوریتم بر روی مجموعه‌ای از تصاویر و قطعات ویدئویی ارائه می‌گردد. در خاتمه نیز به جمع‌بندی و نتیجه‌گیری این تحقیق و تحقیقات آتی در این زمینه خواهیم پرداخت.

فصل ۲- مروری بر فعالیت‌های انجام شده

تا کنون روش‌های زیادی برای تعیین مکان چشم‌ها ارائه شده‌است که در حالت کلی می‌توان آنها را در سه دسته روش‌های مبتنی بر ویژگی^۱، روش‌های مبتنی بر تطبیق الگو^۲ و روش‌های مبتنی بر ظاهر^۳ تقسیم‌بندی نمود.

۱. **روش‌های مبتنی بر ویژگی:** این روش‌ها به جستجوی ویژگی‌های منحصر به فرد در چشم

همچون لبه‌ها، شدت روشنایی عنبیه و توزیع رنگ صلبیه در تصویر می‌پردازند. هدف از این جستجو، شناسایی ویژگی‌های متمایز کننده در اطراف چشم و سپس کلاسه‌بندی نواحی تصویر بر این اساس می‌باشد. اما در صورتی که چشم‌ها بسته باشند یا قسمتی از آنها با مو پوشانده شده باشد یا در اثر چرخش سر قابل رؤیت نباشند، این روش‌ها با شکست مواجه می‌شوند.

۲. **روش‌های مبتنی بر الگو:** در روش‌های مبتنی بر الگو در ابتدا یک مدل عام که در برگی‌رنده

نقوش اصلی چشم می‌باشد طراحی می‌شود، سپس با پویش تصویر مکان‌هایی که منطبق با الگوی چشم هستند، استخراج می‌گردند. نتایج حاصل در این روش به می‌زان زیادی متأثر از سازگاری مدل طراحی شده با شرایط تصویری برداری چهره، مقیاس و چرخش چشم‌ها در تصویر است. همچنین پی‌چیدگی محاسباتی زیاد این روش یکی از عواملی است که استفاده از آن را خصوصاً برای کاربردهای زمان واقعی محدود می‌کند.

۳. **روش‌های مبتنی بر ظاهر:** در روش‌های مبتنی بر ظاهر، محل چشم‌ها بر پایه اطلاعات خام

پیکسل‌ها مانند شدت روشنایی در تصویر مشخص می‌شوند. این روش‌ها معمولاً نیازی به حجم زیادی از داده‌های آموزشی که مربوط به تصاویر چشم در حالات مختلف چرخش سر و شرایط نورپردازی می‌باشد دارند. از این اطلاعات برای آموزش یک کلاسه‌بند مثل شبکه عصبی یا ماشین بردار پشتیبان^۴ استفاده می‌شود. خروجی کلاسه‌بند مکان چشم‌ها را مشخص می‌کند.

¹ Feature based methods

² Template based methods

³ Appearance based methods

⁴ Support vector machine

بعضی از الگوریتم‌های تشخیص چشم ترکیبی از روش‌های قرار گرفته در ۲ یا ۳ گروه را مورد استفاده قرار می‌دهند.

۲-۱- روش‌های تعیین مکان چشم‌ها در تصاویر دو بعدی

۲-۱-۱- روش‌های مبتنی بر ویژگی

Feng و همکارش [۶] با استفاده از تابع واریانس^۱ روشی را برای تعیین موقعیت ویژگی‌های بارز چشم ارائه نمودند. در این روش مدل چشم شامل سه جزء یعنی عنبیه، پلک بالا و پلک پایین می‌باشد. همچنین آنها فرض نمودند که مرز بالا و پایین عنبیه با رئوس پلک‌ها یکدیگر را قطع می‌کنند. بنابراین این مدل با ۶ ویژگی بارز در ناحیه چشم که گوشه‌های چشم، مرزهای بالا، پایین، چپ و راست عنبیه هستند ساخته می‌شود. در این روش تغییر تابع واریانس در جهت افقی برای تعیین مؤلفه عمودی مختصات گوشه‌های چشم و مرزهای چپ و راست عنبیه مورد استفاده قرار می‌گیرد سپس مکان صحیح گوشه‌های چشم و مرزهای چپ و راست عنبیه از بین پیکسل‌هایی که در ناحیه چشم بر روی محور مختصات افقی، ماکزیمم گرادیان را دارند انتخاب می‌شوند. بطور مشابه مرزهای بالا و پایین عنبیه یا پلک‌ها تعیین می‌شوند و عنبیه توسط نقاط مرزی‌اش بازسازی می‌شود. پلک‌ها نیز سهمی‌هایی در نظر گرفته می‌شوند که بکمک سه نقطه، گوشه‌های راست و چپ چشم و رأس سهمی‌ها، ساخته می‌شوند. الگوریتم پیشنهادی بوسیله تصاویری با ابعاد 48×30 از چشم افراد مورد آزمایش قرار گرفت و نتایج خوبی حاصل گردید، البته نرخ آشکارسازی در این آزمایش گزارش نشده‌است و زمان صرف شده برای اجرای الگوریتم روی یک فریم برای یک پردازنده پنتیوم 166MHZ، ۰٫۴ ثانیه گزارش شده است.

Gu و همکارانش [۷] یک آشکارساز لبه SUSAN و گوشه به منظور تشخیص نقاط مربوط به ویژگی‌های چهره پیشنهاد نمودند. در این تحقیق الگوریتمی برای استخراج دو مردمک، گوشه‌های چشم، نقاط بین سوراخ‌های بینی و گوشه‌های دهان ارائه شده‌است. ناحیه چشم‌ها با استفاده از تابع نگاشت انتگرال عمودی^۲ استخراج شده و سپس یک آشکارساز لبه SUSAN بر روی این تصویر اعمال می‌شود. در مرحله بعد عنبیه و پلک بالا خارج از تصویر لبه قطعه‌بندی می‌شوند. در این روش با در نظر گرفتن اینکه گوشه‌های چشم منحنی‌های مرزی احاطه کننده پلک‌ها و عنبیه هستند محل گوشه‌های چشم نیز مشخص می‌شود. الگوریتم برای تصاویر چهره تهیه شده از روبرو، تصاویر چرخانده شده در زوایای بزرگ و کوچک روی دو ضلع راست و چپ تصویر و همچنین تصاویری مربوط به شرایط مختلف نورپردازی و حالات چهره مورد آزمایش قرار گرفت. ارزیابی با مشخص کردن ویژگی‌های چهره روی تصویر بصورت دستی انجام گرفت، نتایج بدست آمده نشان داد الگوریتم پیشنهادی قادر است مکان ویژگی‌های چشم را با خطای پنج پیکسل بدست آورد. رویهم رفته صحت نتایج برای تمام ویژگی‌های چهره ۹۵٪ گزارش شده‌است. گوشه‌های داخلی برای چشم چپ در ۹۸٫۵۲٪ و برای چشم راست در ۹۵٫۹۳٪ موارد درست تشخیص داده شد و نتیجه آشکارسازی گوشه خارجی برای دو چشم نیز ۹۷٫۴۱٪ بیان شده‌است. نتایج مذکور مربوط به تصاویر گرفته شده از روبروی

¹ Variance projection functions

² Variance projection functions

چهره می‌باشند. برای تصاویر چهره همراه با چرخش کوچک سر (کمتر از ۱۵٪) آشکارسازی حدود ۸۵٪ گزارش شده‌است و الگوریتم برای تصاویری با چرخش بیشتر نتیجه قانع کننده‌ای ندارد. همچنین این روش در نور شدیدی که از گوشه‌ها به تصویر تابانیده می‌شود شکست می‌خورد و به مواردی که مربوط به چشم بسته باشد چه در مرحله فرضیات و چه در مرحله آزمایش اشاره نشده است. زمان اجرای این روش برای استخراج تمام ویژگی‌های چهره در تصویری با ابعاد 386×480 حدود ۲۵۶ میلی ثانیه گزارش شده‌است. مساحت ناحیه چشم برای این تصاویر 80×40 پیکسل است. این روش قادر به تشخیص پلک‌ها نمی‌باشد، اما بصورت بسیار تقریبی پلک بالا (برای یافتن گوشه‌های چشم) استخراج می‌شود و هیچ اشاره‌ای نیز به استخراج ویژگی از پلک پایین نشده‌است.

Zhou و همکارش [۸] روشی با استفاده از توابع نگاشت^۱ برای تشخیص چشم پیشنهاد می‌کنند. در این روش تابع نگاشت عمومی^۲ (GPF) بصورت ترکیب از تابع نگاشت انتگرال^۳ (IPF) و تابع نگاشت واریانس^۴ (VPF) بصورت زیر تعریف می‌گردد:

$$IPF'_v(x) = \frac{1}{y_2 - y_1} \int_{y_1}^{y_2} I(x, y) dy \quad 1-2$$

$$IPF'_h(y) = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} I(x, y) dx$$

$$VPF_v(x) = \frac{1}{y_2 - y_1} \sum_{y_i=y_1}^{y_2} [I(x, y_i) - IPF'_v(x)] \quad 2-2$$

$$VPF_h(y) = \frac{1}{x_2 - x_1} \sum_{x_i=x_1}^{x_2} [I(x, y) - IPF'_h(y)]$$

$$\begin{aligned} GPF_v(x) &= (1 - \alpha)IPF'_v(x) + \alpha VPF_v(x) \\ GPF_h(y) &= (1 - \alpha)IPF'_h(y) + \alpha VPF_h(y) \end{aligned} \quad 3-2$$

$$0 \leq \alpha \leq 1$$

در این روابط $I(x, y)$ ، روشنایی در محل (x, y) از تصویر ورودی است. اندیس‌های v, h به ترتیب نگاشت‌های افقی و عمودی در این توابع را نشان می‌دهند.

تابع نگاشت هیبرید^۵ (HPF)، حالت خاصی از تابع نگاشت عمومی می‌باشد که با قراردادن $\alpha = 0.6$ در تابع نگاشت عمومی بدست می‌آید، این مقدار توسط آزمایشات انجام شده تعیین شده‌است. نتایج گزارش شده در این مقاله نشان می‌دهد که تشخیص بر مبنای استفاده از تابع HPF نتایج بهتری نسبت به نتایج

¹ Projection functions

² Generalized projection function (GPF)

³ Integral projection function (IPF)

⁴ Variance projection function (VPF)

⁵ Hybrid projection function (HPF)

توابع VPF و IPF بطور جداگانه دارد. جدول ۱-۲ مقایسه نتایج حاصل شده توسط این سه روش را بر روی سه پایگاه داده مختلف نشان می‌دهد.

جدول ۱-۲: نتایج بدست آمده در روش Zhou و همکارش [۸]

Func	BioID(%)	JAFFE(%)	NJUFace(%)
IPF	۹۳،۶۹	۹۶،۷۱	۹۲،۴۸
VPF	۹۴،۴۱	۹۷،۱۸	۹۵،۵۴
HPF	۹۴،۸۱	۹۷،۱۸	۹۵،۸۲

روشی توسط Tao و همکارش [۹] ارائه گردید که مربوط به تشخیص چشم در تصاویر با پس زمینه‌های شلوغ می‌باشد. در این روش ابتدا با توجه به رابطه هندسی موجود بین اجزای چهره معیار مناسبی برای مکان چشم‌ها تعیین می‌شود. سپس با استفاده از چند آستانه مناسب نواحی کاندید چشم‌ها از تصویر چهره جدا می‌شوند. با توجه به اینکه سطح خاکستری عنبیه و مردمک نسبت به پوست ناحیه اطراف چشم و سفیدی چشم پایین‌تر است بنابراین با یک آستانه مناسب می‌توان چشم را از دیگر اجزای چهره جدا کرد. آستانه بهینه در رنج T_0 تا T_{max} تغییر می‌کند. در این روش T_0 سطح خاکستری متناظر با اولین پیک در هیستوگرام روشنایی تصویر می‌باشد. اما بدلیل نویزی بودن هیستوگرام فوق ممکن است پاسخ الگوریتم به نتایج خطایی منجر شود. برای رفع اثر نویز، هیستوگرام توسط یک فیلتر گوسی هموار می‌گردد [۹].

آستانه T_0 با استفاده از هیستوگرام هموار شده تعیین می‌شود و برای آستانه T_{max} نیز مقدار 0.6 در نظر گرفته می‌شود. تصویر چهره با توجه به آستانه‌های بدست آمده برای T_0, T_{max} باینری می‌شود. زوج بلوک‌هایی که معیارهای مربوط به مکان چشم‌ها را در تصویر برآورده می‌سازند بعنوان کاندید برای چشم راست و چپ تعیین می‌شوند. در نهایت به منظور حذف کاندیدهای نادرست ناحیه متناظر با زوج بلوک‌های کاندید از تصویر سطح خاکستری جدا شده و همبستگی متقابل بین آنها محاسبه می‌شود در صورتیکه مقدار همبستگی متقابل بین آنها بزرگتر از 0.5 باشد دو ناحیه چشم هستند و در غیر اینصورت جستجو برای یافتن چشم‌ها ادامه پیدا می‌کند.

روشی که توسط نصیری و همکارانش [۱۰] ارائه شد به تشخیص چشم در تصاویر رنگی می‌پردازد. الگوریتم مورد نظر با استفاده از سیستم رنگ YCbCr دو نگاشت ی‌کی از روی بخش روشنایی^۱ (EyeMapL) و دی‌گری از روی بخش رنگ^۲ (EyeMapC) می‌سازد. سپس این دو نگاشت با هم ترکیب شده و یک نگاشت نهایی را به وجود می‌آورند. در این روش می‌بایست چهره از روبرو تصویر برداری شود و پوشانیدن چهره توسط عینک و یا هر شیء دی‌گری باعث خطا می‌شود. ای‌ده تولید EyeMapC بر

¹ Luminance

² Chrominance

این اساس است که ناحیه چشم دارای مقدار مؤلفه آبی (Cb) بالا و مؤلفه قرمز (Cr) پایینی می باشد. بعد از اعمال این نگاشت در تصویر سطح خاکستری، محل چشمها از روشهای بالایی برخوردار خواهند بود. در پایان با اعمال یک مرحله پس پردازشی با عملگرهای مورفولوژی و آستانه گذاری محل چشمها مشخص می گردد. روش پیشنهادی برای افزایش دقت از آستانه گذاری متغیر و سپس تست هندسی استفاده می کند، به این معنی که در فرایندی تکراری به ازای چندین آستانه مختلف، نواحی کاندیدی باقی مانده از لحاظ ویژگیهای هندسی منسوب به چشمها بررسی شده و در صورت احراز همه این شرایط، آن جفت به عنوان محل چشمها انتخاب می گردد. برخی از ویژگیهای استفاده شده عبارتند از: فاصله بین دو ناحیه کاندیدی، فاصله بین هر جفت ناحیه از مرکز چهره و موقعیت نواحی در تصویر (چشمها نمی توانند در یک سمت چهره واقع شوند). نرخ کلی تشخیص در این روش حدود ۹۷/۵۷ درصد گزارش شده است.

شاهجویی [۱۱] روشی پیشنهاد داده است که در آن از الگوریتم تجزیه درختی برای تقسیم بندی تصویر چهره استفاده می شود. روش پیشنهادی دارای سه بخش اساسی می باشد: ۱- تقسیم تصویر به بلوکهای جزئی، ۲- ادغام بلوکهای مرتبط و ۳- آشکارسازی چشمها. در اولین بخش که از روش تجزیه درختی مربعی^۱ استفاده می کند، یک تصویر مربعی به چهار بلوک تقسیم می شود و سپس برای هر بلوک معیار همگنی آزمایش می گردد. در صورتیکه یک بلوک معیار همگنی را برآورده سازد، دیگر تقسیم نمی شود در غیر اینصورت دوباره به چهار بلوک تقسیم شده و این مراحل تا زمانی که معیار برای همه بلوکها آزمایش شود، تکرار می گردد. سپس با استفاده از سطح آستانه مناسب که به صورت تجربی محاسبه می گردد بلوکهای ایجاد شده در مرحله قبل ادغام می شوند. در مرحله بعد تمامی پیکسلهایی که شرط سطح آستانه را داشته باشند، برچسب می خورند و پیکسلهایی که تفاوت روشنایی شان با روشنایی پیکسلهای مجاور بیشتر از سطح آستانه باشد دست نخورده باقی می ماند. با توجه به مطالب بیان شده، به دلیل تفاوت زیادی که بین روشنایی سفیدی چشم و مردمک وجود دارد، این نواحی برچسب نمی گیرند. سپس با استفاده از عملگر مورفولوژیکی بستن^۲ مرزهای صورت که برچسب نخورده و به صورت ناپیوسته هستند به هم می پیوندند. از بین این نواحی باقی مانده، ناحیه ای که شروط زیر را برآورده کند به عنوان ناحیه چشم در نظر گرفته می شود:

۱- ناحیه مورد نظر از باقی نواحی بزرگتر باشد.

۲- طول و عرض ناحیه انتخاب شده از اندازه ای که به صورت تجربی به دست آمده، تجاوز نکند.

۳- کل ناحیه در نیمه بالایی تصویر قرار داشته باشد.

در پایان برای استخراج خطوط چشم از یک آشکارساز لبه^۳، با سطح آستانه مناسب استفاده می گردد [۱۱].

¹ Quad Tree Decomposition

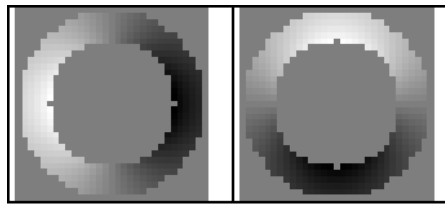
² Closing

³ Edge detector

۲-۱-۲ - روش‌های مبتنی بر تطبیق الگو

Paradas [۱۲] روش دیگری برای استخراج و ردیابی پلک‌ها ارائه نموده‌است. در این روش ناحیه چشم براساس روابط هندسی حاکم بر اجزاء چهره تعیین می‌شود. به منظور تشخیص پلک‌ها، آنها از روشی براساس یافتن قطعه کوچکی در گراف، در صورتیکه گوشه‌های چشم مشخص شده باشند، استفاده کرده‌اند. مدل چشم شامل دو منحنی، یکی برای پلک پایین با یک مینیمم و یکی برای پلک بالا با یک ماکزیمم می‌باشد. در الگوریتم پیشنهادی گوشه‌های چشم با استفاده از یک الگوی پاره‌خطی انعطاف‌پذیر مشخص می‌شوند. الگوی بکاررفته شامل سه پاره‌خط می‌باشد که هیچ‌کدام از آنها به‌تنهایی نمی‌توانند تغییر کنند اما ترکیب آنها برای منطبق شدن بر خطوط اصلی چشم تغییر می‌کند. برای هر جفت از نقاط کاندید برای گوشه چشم الگوریتمی برای جستجوی دو قطعه کوچک بین گوشه‌های چشم مورد استفاده قرار می‌گیرد. با ترکیب و منسجم ساختن دانش قبلی در رابطه با پلک‌ها، تشخیص پلک انجام می‌شود. مثلاً پلک پایین به اندازه مشخصی نسبت به پلک بالا نزول می‌کند. یک جفت از نقاط کاندید که تابع هزینه انتخاب قطعه آنها کمترین مقدار را دارد بعنوان گوشه چشم و قطعه مربوط به آنها بعنوان پلک تعیین می‌شوند. در رابطه با نوع شرایطی که الگوریتم در آن آزمایش شده است یا زمانی که سپری می‌شود گزارشی ارائه نشده است. اما به نظر می‌رسد جستجو برای یافتن قطعه مینیمم بین تمام نقاط کاندید برای گوشه چشم، پروسه زمان‌گیری باشد.

Orazio و همکارانش [۱۳] روشی ارائه دادند که با استفاده از ویژگی‌های هندسی منسوب به عنبیه، نواحی کاندید برای چشم را در تصویر انتخاب می‌کند آنها با استفاده از تقارن و شباهت بین دو چشم، چشم‌ها را در بین نواحی کاندید جستجو می‌کنند. در این روش از یک الگو به شکل دایره برای تشخیص نواحی محتمل بر عنبیه استفاده شده‌است. بدلیل اینکه اندازه عنبیه با توجه به اندازه چشم افراد و فاصله فرد از دوربین متغیر است، یک محدوده برای شعاع این دایره $[R_{\min}, R_{\max}]$ در نظر گرفته می‌شود. در این الگوریتم از روش تبدیل هاف دایره جهت‌دار^۱ برای تشخیص اجزای دایره‌ای در تصاویر استفاده شده‌است. ماسک‌های نشان داده شده در شکل ۱-۲ در هر نقطه جهت بردار شعاعی که با فاصله از مرکز حلقه‌ای با شعاع مینیمم R_{\min} و شعاع ماکزیمم R_{\max} تغییر اندازه داده است را ارائه می‌کنند.



شکل ۱-۲: الگوی چشم در روش Orazio و همکارانش [۱۳]

در ابتدا کانولوشن ماسک و گرادیان تصویر ورودی محاسبه می‌شود، نقاطی که نتیجه کانولوشن برای آنها ماکزیمم مقدار را نشان می‌دهد نقاط محتمل بر وجود عنبیه هستند. سپس برای تعیین میزان شباهت

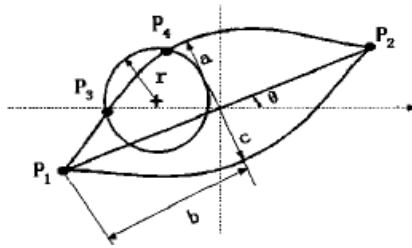
¹ Directional circle Hough Transform

بین نواحی کاندید، متوسط خطای مطلق^۱ بین آنها (رابطه ۲-۴) محاسبه می‌شود و در نهایت با مقایسه مقدار محاسبه شده و یک آستانه مناسب چشم‌ها در تصویر مشخص می‌شوند.

$$MAE_s = \frac{\sum_{i=1}^N \sum_{j=1}^M |a_{ij} - b_{i(M-j+1)}|}{N.M} \quad ۴-۲$$

نتایج در تعیین میزان شباهت وابستگی زیادی به ابعاد ناحیه دارند، همچنین ترکیب‌هایی از دو ناحیه دایره‌ای مشابه در نواحی دیگر چهره بجز چشم نظیر حلقه‌های مو در موهای فردار، گوشه‌های دهان و ... وجود دارد که نتایج را تحت تأثیر قرار می‌دهند. این الگوریتم بر روی تصاویر ۶ فرد در حالات و زوایای مختلف که در دو مجموعه هر کدام شامل ۱۴۲۳ تصویر، تهیه شده بود آزمایش گردید. یک مجموعه شامل تصاویری با چشم‌های کاملاً باز و مجموعه دیگر شامل تصاویری با چشم‌های نسبتاً باز بودند. نرخ آشکارسازی برای تصاویر با چشم باز ۹۶٪ و برای تصاویر با چشم نسبتاً بسته ۴۵٪ گزارش شده است [۱۳].

روش ارائه شده توسط Tan و همکارانش [۱۴]، با استفاده از یک الگوی دگردیس پذیر چشم را از تصویر چهره جدا می‌کند. در این روش از تابع انرژی معرفی شده توسط Tomasi و همکارش [۱۵] جهت ردیابی چشم‌ها استفاده می‌شود. تابع انرژی پیشنهاد شده توانایی ردیابی گوشه‌های چشم را دارد و باعث جذب الگوی چشم به مکان صحیح چشم در تصویر می‌شود. به منظور تخمین گوشه‌های چشم، الگوی مورد استفاده در روش پیشنهادی بصورت پارامتری بیان می‌شود. یک دایره به مرکز X_c و شعاع r و سهمی‌هایی به مرکز X_e که طول آنها با b مشخص می‌شود. ارتفاع منحنی‌های سهمی بالا و پایین با a و c بیان می‌شود. این منحنی‌ها در چهار نقطه P_1, P_2, P_3, P_4 یکدیگر را قطع می‌کنند. بنابراین الگوی چشم‌ها با پارامترهای $(X_c, X_e, r, a, b, c, \theta)$ مشخص می‌شود (شکل ۲-۲).



شکل ۲-۲: الگوی چشم در روش Tan و همکارانش [۱۴]

در این الگو تمام پارامترها اجازه تغییر دارند. معمولاً برای ردیابی و بدست آوردن اطلاعات مربوط به حرکت فقط نواحی از تصویر که بافت غنی دارند مورد استفاده قرار می‌گیرند. اگر یک دنباله تصاویر با $I(X,t)$ نشان داده شود که $X = [u, v]^T$ مختصات یک نقطه از تصویر باشد، در صورتیکه فرکانس نمونه برداری در زمان به اندازه کافی بالا باشد می‌توان فرض کرد نواحی تصویر بدون اینکه در سطح روشنایی آنها تغییری ایجاد شود جابجا می‌شوند. این جابجایی بصورت $I(X, t + \tau) = I(X - d, t)$ بیان می‌شود که d در این رابطه بردار

¹ Mean absolute Error

جابجایی می‌باشد. زمانیکه جابجایی کوچک باشد تابع توصیف کننده سطح روشنایی را می‌توان با سری تیلور تخمین زد:

$$I(X-d, t) = I(X, t) - gd \quad 5-2$$

وظیفه ردیاب یافتن بردار جابجایی d به منظور مینیمم کردن خطا می‌باشد. خطا با رابطه ۶-۲ بیان می‌شود:

$$e = \iint_W [I(X, t) - gd - I(X, t + \tau) - n(X)]^2 \omega dX = \iint_W [h - gd]^2 \omega dX \quad 6-2$$

در این رابطه $h = I(X, t) - I(X, t + \tau) - n(x)$ است و $n(x)$ بیانگر نویز و ω تابع وزن است. با مرتب کردن رابطه بر حسب d و برابر قرار دادن نتیجه با صفر رابطه ۶-۲ بصورت زیر بازنویسی می‌شود:

$$\iint_W [h - gd] g \omega dA = 0 \quad 7-2$$

در رابطه ۷-۲، مساحت پنجره می‌باشد. با توجه به اینکه $(gd)g = (gg)^T d$ و d داخل W ثابت فرض شود تساوی زیر برقرار است:

$$\iint_W hg \omega dA = \left(\iint_W gg^T \omega dA \right) d \quad 8-2$$

در این رابطه $\iint_W gg^T \omega dA$ ، G نامیده می‌شود که یک ماتریس متقارن 2×2 می‌باشد. این ماتریس برای یک پنجره از دو فریم متوالی محاسبه شده، در صورتیکه مقادیر ویژه ماتریس G بزرگتر از یک آستانه مشخص باشند به این معنی است که این پنجره می‌تواند بطور موفقیت آمیزی ردیابی شود. عناصر ماتریس G در رابطه ۹-۲ بیان شده‌اند، در این رابطه، I_x, I_y مشتقات جزئی در جهت‌های X, Y هستند.

$$G = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} \quad 9-2$$

در روش پیشنهادی، تابع وزن برای تمام پیکسل‌ها یک در نظر گرفته می‌شود. در ابتدا تصویر ورودی با الگوی تعریف شده کانالو می‌شود. لبه‌ها، پیک‌ها و دره‌ها در تصویر ورودی با استفاده از عملگرهای مورفولوژیکی استخراج شده و پنجره‌هایی که براساس مطالب گفته شده نواحی مناسب برای ردیابی هستند مشخص می‌شوند. البته قبل از این مراحل با هموارسازی، نویز از تصویر ورودی حذف می‌شود. سپس تابع انرژی، الگوی تعریف شده را به مکان صحیح هدایت می‌کند. برای اینکه تطبیق الگوی تعیین شده با چشم بنحو مؤثرتری انجام شود یک فرایند مینیمم سازی سه مرحله‌ای اجرا می‌شود. مراحل مینیمم سازی به ترتیب زیر می‌باشند:

۱- آشکارسازی عنبیه انجام می‌شود. در این مرحله (X_e, t) به روز می‌شوند.

۲- پلک‌ها در مکان صحیح تنظیم می‌شوند. در این مرحله محدودیت‌های مربوط به توانایی

ردیابی تحت تأثیر تصویر ورودی قرار می‌گیرند، پارامترهای (X_c, a, b, c, θ) تنظیم

می‌شوند و پارامترهای مربوط به عنبیه بدون تغییر باقی می‌مانند.

۳- کلیه پارامترها با در نظر گرفتن تمام جملات انرژی تنظیم می‌شوند.

۲-۱-۳- روش‌های مبتنی بر ظاهر

Jin و همکارانش [۱۶] روشی بر اساس ویژگی‌های ظاهری چشم (روشنایی تصویر) برای تشخیص چشم در تصاویر ارائه دادند. در روش پیشنهادی با استفاده از جمع‌آوری ویژگی‌ها از روی تصاویر یک چشم و تصاویر جفت چشم یک کلاسه‌بند هی‌بری^۱ را آموزش می‌دهند. یک کلاسه‌بند ماکزیمم شباهت^۱ ML و یک کلاسه‌بند ماشین بردار پشتیبان (SVM)^۲ با توانایی دسته‌بندی تصاویر در دو کلاس چشم و غیرچشم آموزش می‌بینند. این دو کلاسه‌بند بصورت متوالی قرار می‌گیرند و در اولین مرحله به تشخیص سریع کاندیدهای چشم می‌پردازند. در مرحله دوم با اعمال عملگرهای مرفولوژی بر روی تصویر باینری که از تصاویر استخراج شده توسط کلاسه‌بندها حاصل شده‌است، خطوط اتصالی باریک و کوتاه حذف می‌شوند. سپس تمام اجزای بهم‌پیوسته باقیمانده برچسب گذاری شده و با مقایسه ابعاد اجزایی که در همسایگی هم هستند، جزئی که ابعاد آن نسبت به همسایگان‌شان خیلی کوچکتر باشد حذف می‌شود. با اعمال یک آستانه مناسب روی ابعاد، اجزای بهم‌پیوسته باقیمانده که کاندیدهای نامناسبی بشمار می‌روند نیز حذف می‌شوند. از طرفی از محدودیت فاصله بین دو چشم برای حذف کاندیدهای باقیمانده نامناسب استفاده می‌شود. فاصله اقلیدسی بین منطقه‌ای که بعنوان کاندید چشم راست و چپ تعیین شده‌است و یک تصویر مرجع از چشم راست و چپ محاسبه می‌شود و کاندیدها براساس این مقدار رتبه‌بندی می‌شوند و بهترین کاندید بعنوان چشم شناخته می‌شود. روش پیشنهاد شده روی چند بانک اطلاعاتی شامل ۵۹۰۲ تصویر چهره در شرایط مختلف نورپردازی، حالات چهره، چرخش سر و پوشانده شدن چهره با عوامل خارجی آزمایش شد که در ۹۰٫۷۵٪ تصاویر چشم تشخیص داده شد. زمان صرف شده برای اجرای الگوریتم با پردازنده پنتیوم ۴، ۴ GHZ و ۲٫۹۳ M و ۲۵۶ M، 105.275 ± 18.619 میلی‌ثانیه گزارش شده است.

Niu و همکارانش [۱۷] روشی پیشنهاد کردند که از کلاسه‌بند Adaboost با ساختار دو بعدی برای تشخیص چشم استفاده می‌کند. منظور از ساختار دوبعدی ترکیب کلاسه‌بندها بصورت متوالی در دو جهت افقی و عمودی می‌باشد.

در ابتدا به منظور آموزش کلاسه‌بندها، از ۳۰٫۰۰۰ تصویر چهره در وضعیت‌های گوناگون به منظور تهیه مجموعه‌ای از نمونه‌های مثبت (تصاویر چشم که در این روش فقط از تصاویر چشم چپ استفاده شده‌است) و مجموعه‌ای از نمونه‌های منفی (که مربوط به نواحی دیگر چهره بجز چشم می‌باشند) بهره‌گیری می‌شود. با تغییر ابعاد و چرخاندن تصاویر چشم تعداد نمونه‌های آموزشی به ۱۵۰٫۰۰۰ نمونه افزایش می‌یابد. در

¹ Maximum Likelihood

² Support Vector Machine

اولین مرحله از پروسه آموزش تعداد ۱۰,۰۰۰ نمونه مثبت بصورت تصادفی از داده‌های آموزشی جدا می‌شوند و با استفاده از نمونه‌های انتخاب شده و تمام نمونه‌های منفی، کلاسه‌بند C_1 آموزش می‌بیند. نرخ آشکارسازی بدست آمده در این مرحله بسیار کم (۵۰٪) می‌باشد. سپس تمام نمونه‌های موجود در گروه چشم توسط کلاسه‌بند C_1 مورد آزمایش قرار گرفتند و در ادامه ۱۰,۰۰۰ تصویر از ۷۵۰۰۰ تصویری که درست کلاسه‌بندی نشده بودند برای آموزش کلاسه‌بند C_2 مورد استفاده قرار گرفتند. پس از n مرحله تکرار الگوریتم، آموزش کلاسه‌بند به پایان می‌رسد (n تعداد ستون‌ها در ساختار دوبعدی کلاسه‌بند می‌باشد). در پروسه تشخیص چشم‌ها در ابتدا ناحیه‌ی چهره توسط یک آشکارساز چهره از تصویر جدا می‌شود و در ادامه از کلاسه‌بند آموزش دیده شده برای تعیین مناطق کاندید برای چشم استفاده می‌شود. در صورتیکه تعداد کاندیدهای انتخاب شده توسط C_i با N_i بیان شود و مرکز p آمین کاندید $C(p)$ باشد مکان چشم L_i با رابطه ۱۰-۲ تعیین می‌شود:

$$L_i = \arg \max_{1 \leq p \leq N_i} \sum_{q=1}^{N_i} \varphi(p, q) \quad 10-2$$

$$\text{where } \varphi(p, q) = \begin{cases} 1 & \text{if } \|c(p) - c(q)\| < \varepsilon \\ 0 & \text{else} \end{cases}$$

در آخرین مرحله از روش پیشنهادی Niu و همکارانش [۲۷] بجای استفاده از ترکیب وزن دار کلاسه‌بندها، به منظور تشخیص دقیق تر از ترکیب کلاسه‌بندها بصورت متوالی استفاده می‌کنند، که مزیت آن کاهش زمان محاسبه بیان شده است.

Xingming و همکارش [۱۸] الگوریتمی برای تشخیص چشم در تصاویر دوبعدی ارائه دادند که به گفته آنها مستقل از شدت روشنایی می‌باشد. در ابتدا با نرمالیزه کردن روشنایی تصویر ورودی، اثر تغییرات سطح روشنایی در آن از بین می‌رود. مراحل نرمالیزه سازی به ترتیب عبارتند از: ۱- تراز کردن هیستوگرام تصویر ورودی، ۲- استفاده از مدل تبدیل affine روشنایی (AT)^۱ [۱۹]. ۳- استفاده از الگوریتم ICR [۱۹] و سپس الگوریتم AT ۴- محاسبه میانگین بین تصاویر حاصل شده در مراحل ۲ و ۳. در روش پیشنهادی فرایند تشخیص چشم‌ها در تصویر ورودی با تشخیص چهره توسط الگوریتم Adaboost در تصویر ورودی آغاز می‌شود، سپس با طی مراحل زیر ویژگی‌های تصویر استخراج می‌شوند:

۱- حذف پوست با استفاده از یک مدل احتمالاتی از رنگ پوست

۲- حذف پس‌زمینه توسط ماسکی که در ناحیه پوست وجود دارد.

۳- باینری ساختن ویژگی‌های چهره

در مرحله بعدی الگوریتم، نقاطی که با در نظر گرفتن قوانین هندسی حاکم بر چشم‌ها از قبیل اندازه، فاصله بین چشم‌ها و موقعیت چشم‌ها در چهره، کاندیدهای ضعیفی برای چشم هستند حذف می‌شوند. نواحی کاندید برای دهان با توجه به اندازه و موقعیت مکانی لب‌ها تعیین می‌شوند و در نهایت با توجه به

¹ Affine transformation (AT) illumination model

ارتباط هندسی بین چشم‌ها و دهان که بصورت یک مثلث می‌باشد نواحی کاندید برای چشم تعیین می‌شوند. در تعیین کاندیدها از ویژگی ابرو و موازی بودن آن با چشم‌ها استفاده شده‌است. در آخرین مرحله الگوریتم از یک ماشین بردار پشتیبان (SVM) برای تعیین نتیجه نهایی استفاده می‌شود.

Kawaguchi و همکارش [۲۰] با استفاده از اطلاعات روشنایی و لبه به تشخیص عنبیه در تصاویر می‌پردازند. در این روش فرض شده است که تصویر ورودی شامل سر و شانه فرد می‌باشد و عنبیه هر دو چشم در تصویر کاملاً مشخص است. در اولین مرحله از الگوریتم ناحیه چهره از تصویر جدا می‌شود که این عمل در تصاویر سطح خاکستری با استفاده از تصویر لبه و در تصاویر رنگی با استفاده از رنگ پوست انجام می‌شود. در مرحله دوم حفره‌ها یا گودی‌ها در تصویر جدا شده، مشخص می‌شوند. برای این منظور تفاضل تصویر اصلی از تصویری که بعد از اعمال عملگر مرفولوژی بستن^۱ بدست می‌آید محاسبه شده و پیکسل‌هایی که مقدار آنها بزرگتر از یک آستانه تعیین شده باشد بعنوان دره شناخته می‌شوند. در مرحله سوم الگوریتم، با استفاده از الگوی تعریف شده توسط Lin و همکارش [۲۱] تابع ارزش برای تمام پیکسل‌های حفره محاسبه شده و ماکزیمم‌های محلی آن بعنوان نواحی کاندید برای عنبیه تعیین می‌شوند. در مرحله چهارم این الگوریتم ابتدا لبه‌ها با استفاده از آشکارساز لبه Canny که روی تصویر چهره اعمال می‌شود، استخراج شده سپس با استفاده از تبدیل هاف دایره‌ای، مناسب بودن کاندیدهای عنبیه روی تصویر لبه بررسی می‌شوند و در آخرین مرحله تابع ارزش برای هر جفت از کاندیدهای عنبیه محاسبه می‌شود، مقدار این تابع ترکیبی از ارزش عنبیه به‌تنهایی و مقدار همبستگی بین دو کاندید و الگوی تعریف شده می‌باشد. با توجه به مقدار تعیین شده در این مرحله محل نهایی عنبیه در تصویر مشخص می‌شود. این الگوریتم بر روی پایگاه داده‌ای شامل ۱۵۰ چهره نرخ آشکارسازی ۹۴٫۷٪ و بر روی پایگاه داده شامل ۱۲۰ چهره که تصاویر همراه با چرخش سر از آن حذف شده بودند نرخ آشکارسازی ۹۸٫۳٪ را نشان می‌دهد.

Suandi و همکارانش [۲۲] روشی برای تعیین جهت نگاه افراد با استفاده از مردمک و رنگ پوست ارائه می‌کنند. این مقاله روشی جدید در تشخیص وضعیت افقی سر از روی مجموعه تصاویر وی‌دئویی رنگی برای کاربردهای بلادرنک معرفی می‌کند. ایده اصلی در این الگوریتم این است که هنگامی که سر در وضعیت دلخواهی قرار می‌گیرد (به سمت چپ یا راست) ارتباط هندسی کاملاً مشخصی بین فاصله از مرکز دو مردمک و فاصله بین مراکز دو مردمک وجود دارد. از روی این دو فاصله، نسبتی با نام "نسبت افقی" محاسبه می‌شود. این نسبت، در کاهش وابستگی ویژگی‌های چهره به نویز جهت ردیابی دقیق و پایدار مؤثر است، علاوه بر این از این نسبت برای تشخیص زاویه چرخش چهره استفاده می‌شود. روش ارائه شده بسیار ساده و با پیچیدگی محاسباتی کم می‌باشد. این الگوریتم برخلاف سایر روشها که از گوشه‌های داخلی و بی‌خارجی چشم برای تشخیص وضعیت افقی سر استفاده می‌کنند از مردمک برای این هدف بهره می‌گیرد. استفاده از مردمک همچنین می‌تواند به تشخیص جهت و می‌زان باز بودن چشم نیز کمک کند.

¹ Closing

۲-۲- روش‌های ردیابی

Tian و همکارانش [۲۳] روشی برای شناسایی سه وضعیت چشم - کاملاً بسته، نیمه باز و کاملاً باز - با استفاده از موجک گابور^۱ در دنباله‌های تصاویر نزدیک به تمام‌رخ ارائه دادند. ویژگی‌های بکار رفته در این روش، گوشه داخلی، گوشه خارجی و نقطه بین آنهاست. گوشه‌های چشم در تمام دنباله ویدئویی ردیابی می‌شوند. مهمترین آنها برای ردیابی، گوشه داخلی چشم می‌باشد چون مکان بقیه ویژگی‌ها با توجه به مکان گوشه داخلی تعیین می‌شود. مکان اولیه ویژگی‌ها نیز بصورت اطلاعات اولیه مفروض است. ردیابی با مینیمم ساختن تابعی روی یک جابجایی خاص انجام می‌شود. این تابع وابسته به مقادیر سطح روشنایی می‌باشد. گوشه‌های خارجی چشم نیز با استفاده از ابعاد چشم‌ها در فریم اول بدست می‌آیند. برای هر کدام از این سه ویژگی یک مجموعه از ضرایب گابور در ابعاد و زوایای متعدد^۲ محاسبه می‌شود. سه فرکانس فضایی با عدد موج‌های $k_i = \left(\frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{8}\right)$ و شش چرخش از 0 تا π با فاصله‌های $\frac{\pi}{6}$ مورد استفاده قرار می‌گیرد. این ۱۸ ضریب برای تعیین حالت چشم به یک شبکه عصبی وارد می‌شوند. تشخیص حالت چشم در روش ارائه شده موفقیت آمیز گزارش شده است ولی گزارشی از چگونگی ردیابی گوشه‌های چشم وجود ندارد. نرخ شناسایی زمانیکه هر سه وضعیت (چشم باز، بسته و نیمه باز) وجود دارد ۸۹٪ و زمانیکه دو چشم بطور کامل در تصویر وجود دارند ۹۳٪ اعلام شده است.

Sirohey و همکارانش [۲۴] علاوه بر تشخیص چشم‌ها، به ارائه الگوریتمی برای ردیابی پرداخته‌اند. روش آنها برای آشکارسازی برپایه یافتن ترکیبی از لبه‌های جدا شده از تصویر که پلک بالا را توصیف می‌کنند، عمل می‌کند. در ابتدا حرکت سر از روی ترکیب لبه‌ها و طرح سر مشخص می‌شود. براساس این اطلاعات محل سر تثبیت^۳ می‌شود. تفاضل بردار جابجایی سر از تغییر مکان عنیبه و پلک محاسبه می‌شود و تنها اطلاعات مربوط به حرکت مستقل آنها باقی می‌ماند. برای ردیابی پلک‌ها، در ابتدا حرکت لبه در پلک‌ها با توجه به جهت و دامنه بردارهایی که به آنها مرتبط شده است، مشخص می‌شوند. در صورتیکه پیکسل‌های لبه در همسایگی پیکسل‌های علامت‌گذاری شده قرار گیرند متعلق به پلک هستند. سپس پیکسل‌های کاندید شده با چند جمله‌ای درجه سوم برازش^۴ داده می‌شوند. با این روش عنیبه و پلک‌ها در هر فریم بطور صحیح مشخص می‌شوند. در روش پیشنهادی اطلاعات مربوط به پلک‌زدن با توجه به فاصله رأس پلک بالا از مرکز عنیبه تعیین می‌شود. در دنباله تصاویری با ۱۲۰ فریم از یک شخص که شامل تصاویری با عینک و فاقد عینک می‌باشد، ردیابی در ۹۰٪ فریم‌ها با موفقیت انجام شده است. به نتایج مربوط به مدل‌سازی پلک پایین، تشخیص و ردیابی آن اشاره نشده است.

Black و همکارش [۲۵] یک الگوریتم مبتنی بر تطبیق الگو ترکیب شده با شار نوری^۵ ارائه کرده‌اند که حرکت ثابت و تغییرشکل پذیر ویژگی‌های چهره را با استفاده از مدل‌های پارامتری تکه‌ای از حرکت تصویر

¹ Gabor wavelet

² Multi-scale and multi-orientation

³ stabilized

⁴ fitted

⁵ Optical flow

نشان می‌دهد. منظور از ویژگی‌های چهره، صورت، ناحیه چشم‌ها، ابروها و دهان می‌باشد. چهره به چند قسمت تقسیم می‌شود و حرکت هر قسمت بطور مستقل با مدل‌های حرکت برنامه‌ریز^۱ مدل می‌شود. حرکت چشم‌ها بصورت زیر مدل می‌شود:

$$u(x, y) = a_0 + a_1x + a_2y$$

۱۱-۲

$$v(x, y) = a_3 + a_4 + a_5y$$

u, v اجزای افقی و عمودی حرکت در نقطه $p(x, y)$ در تصویر هستند. مختصات با توجه به بعضی از نقاط تصویر (مثلاً مرکز نواحی) تعیین می‌شوند و تفاضل بین تصویر و تغییر پارامترهای بدست آمده در مشاهده قبلی نیز توسط روش کاهش گرادیان^۲ مینیمم می‌گردد. گذر حالت چشم را می‌توان با سه پارامتر انتقال عمودی، اختلاف^۳ و تغییر شکل سریع که در جدول ۲-۲ بیان شده‌است، نشان داد.

جدول ۲-۲: نتایج بدست آمده در روش Black و همکارانش [۲۵]

پارامتر	توصیف ریاضی	آستانه	نتیجه
انتقال	a_3	<-0.1 >0.1	بالا پایین
اختلاف	$a_1 + a_5$	>0.02 <-0.02	بسط انقباض
تغییر شکل	$a_1 - a_5$	>0.005 <-0.005	تغییر شکل افقی تغییر شکل عمودی

با ترسیم منحنی‌های سه پارامتر مذکور برحسب زمان و تعیین مینیمم‌ها و ماکزیمم‌های محلی آن، مشاهده می‌شود که تغییرات در توابع بصورت همزمان اتفاق می‌افتد. پلک زدن زمانی صورت می‌گیرد که شروط ماکزیمم انتقال و ماکزیمم اختلاف و ماکزیمم تغییر شکل، همزمان باهم برقرار شوند. نتایج بدست آمده، صحت و درستی ۸۸٪ را برای دنباله‌های مصنوعی ساخته شده و ۷۳٪ را برای تصاویر سینمایی در حالات مختلف چهره نشان می‌دهند. زمان پردازش هر فریم ۱۲ دقیقه گزارش شده‌است که برای کاربردهای بلادرنگ مناسب نیست.

Moriyama و همکارانش [۲۶] روشی ارائه کرده‌اند که از مدل چشم همراه با جزئیات استفاده کرده‌است. یک الگو با استفاده از دو نمونه پارامتر، ساختار و حرکت، ساخته می‌شود. پارامترهای ظاهری، ساختار ناحیه چشم را با در نظر گرفتن نژاد و تغییرات فردی و سنی توصیف می‌کنند. آنها شامل اندازه، رنگ عنبیه، صلبیه، مناطق تی‌ره در اطراف گوشه‌های راست و چپ چشم، پلک‌ها، طول و ضخامت چین خوردگی پلک‌ها، وسعت برآمدگی زیر چشم، پهنای بازتاب نور روی برآمدگی زیر چشم و گودی زیر آن می‌باشند. پارامترهای مربوط به حرکت در طول زمان تغییر می‌کنند. حرکت عنبیه با مختصات دوبعدی

¹ Planner motion models

² Gradient descent

³ Divergence

مرکز آن بی‌ان می‌شود. ارتفاع پلک‌ها نیز باز و بسته شدن چشم را نشان می‌دهند. انحراف پلک بالا نیز یکی از پارامترهای حرکت می‌باشد که تغیی‌رات پلک بالا را در زمان حرکت مردمک نشان می‌دهد. پارامترهای ساختار بطور خودکار مشخص نمی‌شوند و مدل با تنظیم پارامترهای ساختار بصورت دستی تعیی‌ن می‌شود. برای شروع الگوریتم، پارامترهای ساختار استخراج شده و در تمام فریم‌های دنباله ویدئویی ثابت نگهداشته می‌شوند. علاوه بر آن وی‌ژگی‌ها با تکرار فرای‌ند می‌نی‌مم سازی می‌انگی‌ن مربعات خطا بی‌ن تصویری ورودی و الگوریتمی می‌شوند.

در این روش، مدل ردیابی وی‌ژگی‌های چشم و آشکارسازی پلک زدن قسمتی از یک سیستم برای شناسایی اتوماتیک بود. آنها در روش‌شان فرضیه‌ای را آزمایش کردند که بر اساس وابستگی بی‌ن حرکت سر، جهت چشم و تغییری مکان لب در طول لبخند عمل می‌کرد.

یک روش کانتور فعال توسط Paradas [27] برای ردیابی پلک‌ها پیشنهاد شد. در این روش مدلی برای چشم شامل دو منحنی یکی برای پلک پایینی با یک می‌نی‌مم و یکی برای پلک بالا با یک ماکزیمی‌مم معرفی شد. ردیابی پلک‌ها با روش کانتور فعال انجام می‌شود که حرکت در آن با کوچک کردن انرژی مارها تعیی‌ن می‌گردد. یک مار بسته که پلک را ردیابی می‌کند با انتخاب درصد کوچکی از پی‌کسل‌های اطراف کانتورهای بدست آمده در هنگام ارزش دهی آغازینی یا ردیابی روی فریم‌های قبلی، ساخته می‌شود. گوشه‌های چشم نیز در می‌ان این نقاط می‌باشند. خطای جبران سازی حرکت برای هر مار (x_0, y_0) در می‌ان محدوده مجاز تغییری (dx, dy) محاسبه می‌شود. پی‌کسل‌های $(x_0 + dx, y_0 + dy)$ که کوچکتری‌ن خطا را تولید می‌کنند بعنوان کاندیدی‌دهای ممکن برای مار در فریم جاری می‌باشند. در مقاله ارائه شده گزارشی در رابطه با زمان اجرای الگوریتم داده نشده است. اما به مقاوم بودن آن در مقابل پلک زدن، انتقال و چرخش سر، بزرگی اندازه و محل نمایی‌ان بودن چشم‌ها اشاره شده است.

روشی توسط Garcia [28] و همکارانش ارائه گردید که با استفاده تکنولوژی نور مادون قرمز به تشخیص و ردیابی چشم‌ها می‌پردازد. در این روش تجهیزات سخت‌افزاری مورد استفاده شامل ۲ دوربین، ۴ نورافکن و ۴ کامپیوتر که در شبکه قرار داشتند می‌شد که دوربین اصلی در این سیستم مجهز به یک CCD حساس به نور مادون قرمز بود. این دوربین برای تهیه یک تصویر نمای نزدیک از شخص بکار گرفته شد. در تصویر حاصله ۴ نقطه نورانی وجود داشت که برای تعیین جهت نگاه مورد استفاده قرار گرفت. در پروژه مذکور به منظور تأمین نور مناسب برای این دوربین از نورافکنی بفرم L و متشکل از LEDهای مادون قرمز استفاده شد. یک دوربین دیگر نیز برای آشکارسازی مکان سر در تصویر و بازیابی مکان چشم، در صورتیکه مکان چشم‌ها در فرایند ردیابی از دست برود، مورد استفاده قرار گرفت. به کمک یک واحد کنترل کننده، دوربین بنحوی حرکت افقی-عمودی (pan-tilt) سریع را انجام می‌دهد تا مکان چشم در وسط تصویر قرار گیرد. فرمان‌های کنترل این واحد نیز از طریق پورت RS232 کامپیوتر صادر می‌شوند. از مشخصات روش پیشنهادی وجود ۴ نورافکن به شکل L و استفاده از ۴ کامپیوتر که به منظور کاهش زمان، پروسه عملیات پردازش تصویر را بصورت موازی انجام می‌دهند می‌باشد.

ردیابی در این روش با یک تصویر ورودی در رزولوشن 768×586 کار می‌کند پروسه ردیابی به کمک الگوریتمی که به ۴ فاز تقسیم می‌شود اجرا می‌گردد. در ابتدا مکان مردمک با اندازه‌گیری توده مرکز مردمک تخمین زده می‌شود سپس با یافتن مرزهای عنبیه شعاع آن مشخص می‌شود. در مرحله بعد لکه‌های داخل عنبیه برای تشخیص نقاط نورانی آنالیز می‌شوند و در آخرین مرحله با توجه مرکز مردمک و مکان ۴ نقطه نورانی جهت نگاه فرد تعیین می‌شود.

Fathi و همکارش [۲۹] روشی برای تشخیص و ردیابی چشم در یک دنباله ویدیویی با استفاده از ویژگی رنگ پوست ارائه نمودند. برای تعیین مدل رنگ پوست فضای رنگ RGB به مدل rg نرمالیزه شده منتقل می‌شود. این عمل با تبدیل فضای ۳ بعدی به ۲ بعدی باعث کاهش پیچیدگی می‌شود و حساسیت مقادیر RGB در شرایط متفاوت نوری را نیز کم می‌کند. سپس با استفاده از هیستوگرام مجموعه‌ای از تصاویر که بصورت دستی برچسب‌گذاری شده‌اند و در دو کلاس پوست و غیرپوست حضور دارند، توزیع‌های احتمالاتی گسسته $p(rg|non-skin)$ و $p(rg|skin)$ محاسبه می‌شوند. در ادامه با استفاده از قانون بیز احتمال $p(skin|rg)$ برای پیکسل‌های تصویر ورودی محاسبه می‌شود و در نهایت مقایسه با یک آستانه مناسب، پیکسل را در یکی از دو گروه پوست یا غیرپوست قرار می‌دهد. احتمال پیشین برای دو کلاس نیز از روی هیستوگرام‌ها قابل محاسبه می‌باشد. بعد از کلاسه‌بندی پیکسل‌ها بزرگترین سطح متقارن بعنوان پوست در نظر گرفته می‌شود و سپس ناحیه چهره از تصویر جدا شده و به یک تصویر سطح خاکستری تبدیل می‌شود. در الگوریتم پیشنهادی سه مرحله برای تعیین مناطق کاندید چشم وجود دارد. در ابتدا با توجه به اینکه سطح روشنایی اجزای چهره نسبت به پوست پایین‌تر است، با یک آستانه مناسب این نواحی از چهره جدا می‌شوند و بعنوان کاندیدهای چشم برچسب‌گذاری می‌شوند، در دومین قدم نواحی کاندید بر اساس اندازه و مکان اصلاح می‌شوند و در آخرین مرحله با استفاده از فیلتر واریانس چشم که با رابطه ۲-۱۲ بیان می‌شود، مکان دقیق چشم‌ها تعیین می‌شوند. در روش پیشنهادی واریانس با استفاده از ۴۰ تصویر چشم با ابعاد 33×24 و یک ماسک با ابعاد 3×3 محاسبه می‌شود.

$$F_e(i, j) = \frac{1}{N} \sum_{k=1}^N I_{\delta}^K(i, j) \quad 12-2$$

$I_{\delta}^K = (i, j)$ واریانس (i, j) امین بلوک از k امین تصویر آموزشی می‌باشد.

در مرحله تست، تصویری که بالاترین مقدار را در پاسخ به این فیلتر داشته باشد بعنوان ناحیه دقیق چشم در نظر گرفته می‌شود. در این الگوریتم با تعیین یک ناحیه با ابعاد 40×40 پیکسل بعنوان ناحیه جستجو در اطراف چشم، ناحیه‌ای که در فریم بعدی نزدیکترین فاصله را نسبت به مکان قبلی داشته باشد و همبستگی آن با چشم از یک مقدار آستانه بیشتر باشد بعنوان چشم شناخته می‌شود. در مراحل مربوط به ردیابی بردار واریانس ناحیه چشم برای شناسایی تغییرات در شکل چشم مورد استفاده قرار می‌گیرد.

روش دیگری توسط Haro و همکارانش [۳۰] ارائه شد که با توجه به ویژگی‌های فیزیولوژیکی، حرکتی و ظاهر^۱ به تشخیص چشم می‌پردازد. در این روش از یک منبع نور مادون قرمز و یک دوربین حساس به

¹ Appearance

نور مادون قرمز استفاده شده است. در این روش نیز مشابه با روش‌های دیگری که از نور مادون قرمز استفاده می‌کنند، دو تصویر بدست می‌آید که در یکی از آنها مردمک با سطح روشنایی بسیار بالا و در دیگری مردمک با سطح روشنایی بسیار پایین نشان داده شده است. الگوریتم پیشنهادی از تفاضل این دو تصویر که با وجود نویز شامل اطلاعات مفیدی می‌باشد، برای تشخیص چشم‌ها استفاده می‌کند. در مرحله دوم، از یک آستانه تطبیق‌پذیر^۱ به منظور حذف پیکسل‌هایی که مربوط به چشم نیستند استفاده می‌شود. برای این منظور بعد از محاسبه هیستوگرام این تصویر، پیکسل‌هایی که سطح روشنایی آنها از یک مقدار آستانه پایین‌تر می‌باشند حذف می‌شوند. بعد از حذف اجزای بهم‌پیوسته‌ای با کمتر از سه پیکسل، نواحی باقیمانده بعنوان کاندیدهای حضور مردمک شناخته می‌شوند. سپس یک پنجره با ابعاد 16×16 پیکسل که مرکز آن بر روی روشن‌ترین پیکسل در یک از گروه از پیکسل‌ها قرار می‌گیرد، منطبق شده و با دقت به این نکته که پنجره دو مردمک همپوشانی ندارند تعداد کاندیدها کاهش می‌یابد. در روش پیشنهادی از کالمن فیلتر به منظور ردیابی استفاده شده است. برای ردیابی از بردارهای حالت ۴ بعدی، شامل مکان x ، مکان y ، سرعت در جهت x و سرعت در جهت y استفاده می‌شود. بدلیل اینکه شتاب اطلاعات مفیدی در ردیابی ایجاد نمی‌کند در نظر گرفته نمی‌شود. با استفاده از ماتریس کواریانس کالمن فیلتر میزان شباهت بین حرکت یک منطقه کاندید و حرکت مردمک اندازه‌گیری می‌شود. رابطه ۲-۱۳ معادله مربوط به ماتریس کواریانس تصحیح شده^۲ را نشان می‌دهد.

$$P_k = (1 - k_h H) P'_k \quad 13-2$$

در این رابطه P_k, P'_k به ترتیب ماتریس کواریانس قبل و بعد از زمان k هستند، ماتریس H رابطه بین بردار حالت و بردار اندازه‌گیری و k_h ماتریس بهره کالمن می‌باشد. برای اندازه‌گیری صحت و درستی ردیابی بصورت احتمالاتی رابطه ۲-۱۴ محاسبه می‌شود.

$$P(\bar{x}_{k+1}) = N(\bar{x}_k, P_k) \quad 14-2$$

در صورتیکه نتیجه نزدیک به صفر باشد به این معنی است که عمل ردیابی بخوبی انجام نمی‌شود زیرا موقعیت در مقایسه با کواریانس بین دو فریم بشدت تغییر کرده است و اگر نزدیک به یک باشد به این معنی است نتیجه ردیابی منطقی می‌باشد. رابطه ۲-۱۵ تخمینی از میزان حرکت ناحیه موردنظر را نشان می‌دهد.

$$M = 1 - P(\bar{x}_{k+1}) \quad 15-2$$

در صورتیکه مقدار M نزدیک به صفر باشد نشان دهنده سکون ناحیه و در صورتیکه نزدیک به یک باشد نشان دهنده حرکت ناحیه است. در مرحله بعد یک مدل احتمالاتی مبتنی بر ظاهر^۳ برای چشم تهیه می‌شود و با آنالیز آماری بافت برای نواحی مختلف (چشم و غیرچشم) امکان کلاسه‌بندی فراهم می‌گردد. ابتدا با استفاده از تصاویر آموزشی شامل ۸۵ تصویر چشم و ۱۰۳ تصویر غیرچشم دو فضای برداری برای

¹ Adaptive thresholding

² Update

³ Probabilistic appearance based model

کلاس چشم و کلاس غیرچشم تهیه می‌شود، مردمک در این تصاویر بصورت تیره ظاهر می‌گردد. ترکیب ویژگی‌های بافت و اطلاعات ردیابی چشم‌ها به این صورت مشکل می‌باشد بهمین دلیل در این الگوریتم از آنالیز احتمالاتی اجزای اصلی^۱ (PPCA) استفاده شده‌است. PPCA یک توزیع گوسی را برای داده‌های آموزشی در نظر می‌گیرد. برای کلاسه‌بندی ناحیه^۲ t ، احتمال $p(t)$ برای دو فضای برداری چشم و غیر چشم محاسبه می‌شود و ناحیه^۲ t به فضایی تعلق پیدا می‌کند که احتمال مربوط به آن مقدار بیشتری داشته باشد. تجزیه تحلیل عوامل شبیه به PCA فضای N بعدی داده‌ها را به فضایی با D بعد کاهش می‌دهد. با استفاده از تحلیل عوامل ناحیه^۲ N بعدی t به بردار با نام x با D بعد کاهش می‌یابد (۲-۱۶).

$$t = Wx + \mu + \varepsilon \quad ۱۶-۲$$

در رابطه^{۲-۱۶}، W وزن‌های عاملی^۲ را نشان می‌دهد، μ بردار میانگین داده‌های آموزشی و ε نیز خطا می‌باشد. معمولاً مدل نرمال برای t در نظر گرفته می‌شود. همانطور که گفته شد PPCA احتمال تعلق یک ناحیه را به کلاس چشم یا غیرچشم تعیین می‌کند اما بدلیل اینکه این احتمالات در حالت گذرا ناپایدارند و وابسته به مناطقی هستند که در نزدیکی چشم قرار دارند، برای تصمیم‌گیری کافی نیستند. در روش پیشنهادی، با ترکیب وزن‌دار احتمالات بدست آمده برای ناحیه^۲ t در زمان i ، رابطه^۲ جدیدی بصورت زیر بدست می‌آید که مبنای کلاسه‌بندی یک ناحیه در گروه چشم یا غیرچشم می‌باشد.

$$P_i(t) = \alpha P_{eye}(t) + \beta P_{noteye}(t) + \gamma M + P_{i-1}(t) \quad ۱۷-۲$$

α, β, γ بر اساس نتایج PPCA و ردیاب کالمن تغییر می‌کنند. این ثابت‌ها در ابتدا و در فریم اول، برای تمام نواحی صفر تعریف می‌شوند. در آخرین مرحله باید محل چشم‌ها از بین کاندیدها انتخاب شود به این منظور از یک مدل مبتنی بر ظاهر برای یافتن $1/4$ بالایی چهره استفاده می‌گردد. این مدل با استفاده از ۱۶۵ تصویر، آموزش می‌بیند. در این الگوریتم زوج نواحی که در $1/4$ بالایی تصویر قرار بگیرند بعنوان چشم شناخته می‌شوند. نتایج گزارش شده مربوط به دو وضعیت حرکت آرام و کم سر و حرکت سریع و زیاد سر می‌باشد. نتایج حاصل از روشی که تنها از فیلتر کالمن به منظور ردیابی استفاده می‌نمود نیز در مقایسه با روش پیشنهادی ارائه شده‌اند (جدول ۲-۳).

جدول ۲-۳: نتایج بدست آمده در روش Haro و همکارانش [۳۰]

حرکت سریع و زیاد سر

حرکت آرام و کم سر

Detected eyes (900 frames)	Kalman filtering	Proposed method	Detected eyes (900 frames)	Kalman filtering	Proposed method
Average detected	3.95540	2.10813	Average detected	4.89608	1.80335

¹ Probabilistic Principal Component analysis (PPCA)

² Factor Loadings

زمان صرف شده در این الگوریتم در سیستمی با یک پروسوسور پنتیوم II، 200MHZ، ۲۵ فریم در هر ثانیه و برای سیستمی با دو پروسوسور پنتیوم II، 200MHZ، ۲۹ فریم در هر ثانیه گزارش شده است. رزولوشن تصاویر در این آزمایش برابر با 320×240 می باشد.

Hiley و همکارانش [۳۱] با تغییر در تجهیزات سخت افزاری، هزینه پرداخت شده در روش هایی که از نور مادون قرمز استفاده می کردند را کاهش دادند. در این روش برای تولید ویژگی ها از پنچ LED منتشر کننده نور مادون قرمز استفاده شده است که بوسیله پورت موازی برای تولید نور مناسب کنترل می شوند، چهار LED با زاویه دید بالا در چهار گوشه مونیتور قرار گرفته اند، زمانیکه این LEDها تابش می کنند چهار نقطه نورانی در اثر بازتاب از قرنیه در تصویر بوجود می آید. پنجمین LED به منظور ایجاد نقطه نورانی در تصویر در مقابل لنز قرار می گیرد. یک webcam قادر به تصویربرداری با نور مادون قرمز نمی باشد، در روش پیشنهادی با حذف فیلتر بلوکه کننده نور مادون قرمز و استفاده از فیلتری برای حذف نور مرئی شدید (که باعث ایجاد پدیده اشباع می شود) Webcam تبدیل به دوربینی با قابلیت تصویر برداری با نور مادون قرمز شد.

این روش نیز همانند روش های دیگری که از نور مادون قرمز استفاده می کنند، اثر کنتراست بین مردمک و عنبیه را افزایش می دهد و سپس تفاضل دو تصویر با مردمک روشن و مردمک تیره محاسبه می شود. تصویر تفاضل با استفاده از یک آستانه از پیش تعیین شده باینری می گردد و به این ترتیب ناحیه مردمک از تصویر کاملاً جدا می شود. برای تعیین مرکز مردمک از تبدیل هاف استفاده می شود. در ابتدا شعاع مردمک تخمین زده شده و بر پایه این شعاع دایره های روی مردمک برازش داده می شود. پروسه حدس شعاع و جایگذاری دایره تا جائیکه مرکز مردمک با خطای مشخصی تعیین گردد، ادامه می یابد. این روش به همراه عملگرهای مرفولوژی یک روش مقاوم به نویز می باشد.

الگوریتمی توسط Hansen و همکارش [۳۲] بر مبنای تصویر برداری مادون قرمز با هدف ایجاد دو تصویر مختلف با مردمک تیره و روشن ارائه شد. در این روش اولین قدم تعیین نقاط کاندید برای مردمک می باشد. کاندیدها نقاطی هستند که بالاترین تفاوت در سطح روشنایی را بین تصویر با مردمک تیره و مردمک روشن دارند. با استفاده از کلاسه بند Adaboost [۳۳] تعدادی از کاندیدهای نادرست فیلتر می شوند اما ردیابی و تخمین مکان چشم ها در این روش با استفاده از فیلتر کردن جزئی^۱ انجام می شود. اگر یک دنباله با T فریم داده شده باشد، حالات و اندازه گیری ها در زمان t به ترتیب با y_t, x_t نشان داده می شوند، حالات و اندازه گیری های قبل از آن نیز به ترتیب با $\underline{y}_t = (y_1, \dots, y_t), \underline{x}_t = (x_1, \dots, x_t)$ بیان می شوند. این فیلتر با استفاده از یک مجموعه نمونه های وزن دار $S_t^N = \{(x_t^{(n)}, \pi_t^{(n)})\}_{n=1}^N$ که π آمین نمونه از حالت x در زمان t و با وزن $\pi_t^{(n)}$ می باشد، توزیع $p(x_t | \underline{y}_t)$ را تخمین میزند. این مجموعه از نمونه ها به مجموعه جدید S_{t+1}^N گسترش می یابد که احتمال پسین $p(x_{t+1} | \underline{y}_{t+1})$ را بیان می کند. مکان جزء مورد نظر معمولاً با میانگین نمونه ها ارائه می شود. اما در تصاویر بزرگ نیاز به اجزای زیادی از تصویر به منظور نمونه گرفتن از پارامترهای مکانی می باشد که باعث افزایش در پیچیدگی و زمان محاسبه خواهد شد. سپس الگوریتم

¹ Particle Filtering

جابجایی میانگین^۱ برای یافتن مُد غالب مورد استفاده قرار می‌گیرد. محاسبه نسبت شباهت^۲ برای اثبات فرضیه حضور ناحیه چشم در نقطه کاندید شده استفاده می‌گردد. نسبت شباهت با استفاده از رابطه ۱۸-۲ بدست می‌آید. در این رابطه $\xi(y)$ شباهت بین ناحیه هدف و نواحی کاندید را محاسبه می‌کند (رابطه ۱۹-۲).

$$f_R(\varphi | \mu) = \frac{f_e}{f_a} = \frac{\xi(y)}{\prod_i f_L[\Delta I(i)]} \quad 18-2$$

$$\xi(y) \equiv \sqrt{1 - \rho(y)} \quad , \quad \rho(y) \equiv \rho[p(y), q] = \sum_{z=1}^m \sqrt{p_z(y)q_z} \quad 19-2$$

در این رابطه مدل زمینه را تعریف می‌کند، زمینه شامل پیکسل‌هایی از تصویر است که مربوط به چشم نیستند. توزیع پیکسل‌ها در این مدل با استفاده از لاپلاسین ساخته می‌شود (رابطه ۲۰-۲).

$$f_L(\Delta I) = \frac{1}{Z_L} = \exp\left(-\left|\frac{\Delta I}{\delta}\right|\right) \quad 20-2$$

ΔI تصویری است که از تفاضل تصاویر مربوط به مردمک روشن و مردمک تیره ساخته شده است و δ طول تابع لاپلاسین می‌باشد. توزیع پیکسل‌ها در زمینه بصورت $f_a(I) = \prod_{i \in \lambda_h} f_L[\Delta I(i)]$ می‌باشد. λ_h در این رابطه یک ناحیه فرضی است.

Kim و همکارش [۳۴] روشی برای تشخیص و ردیابی مردمک ارائه داده‌اند که از شبکه‌های عصبی و الگوریتم جابجایی میانگین استفاده می‌کند. در این روش برای غلبه بر مشکل حرکات سریعی سر ابتدا ناحیه صورت با بهره‌گیری از مدل رنگ پوست و آنالیز اجزای بهم‌پیوسته^۳ تشخیص داده می‌شود. سپس ناحیه چشم با استفاده از کلاسه‌بند بافتی که بر پایه شبکه‌های عصبی عمل می‌کند، مشخص می‌گردد. این کلاسه‌بند، نواحی چهره را به دو کلاس چشم و غیر چشم تقسیم بندی می‌کند. پس از این‌که ناحیه چشم مشخص شد، به طور پیوسته و صحیح توسط الگوریتم جابجایی میانگین ردیابی می‌شود. در اولین مرحله فضای رنگی RGB به فضای کروماتیک^۴ تبدیل شده و با توجه به اینکه توزیع رنگ برای تصاویر چهره در این فضا به ناحیه بسیار کوچکی محدود می‌گردد، این تصاویر با یک توزیع گوسی دو بعدی تخمین زده می‌شود. با آموزش مدل و مطابقت تصویر ورودی با آن، چهره در تصویر مشخص می‌گردد. با آستانه گذاری روی تصویر تطبیقی یافته تصویر با نری حاصل می‌شود و سپس با اعمال آنالیز اجزای بهم‌پیوسته نواحی کوچک نویز مانند حذف و محل اصلی چهره مشخص می‌گردد. در مرحله دوم با توجه به ویژگی‌های زیر تشخیص انجام می‌گردد:

¹ Mean Shift

² Likelihood ration

³ Connected Component

⁴ Chromatic color

۱. کنتراست روشنایی بی‌نواحی سفیدی چشم، عنبیه و مردمک همراه با بافت پلک بسیاری از زی‌باد است.

۲. چشم در بخش بالایی ناحیه صورت واقع است.

استفاده از این دو ویژگی بی‌چی‌دگی مسأله را کاهش داده و جداسازی بی‌نواحی چشم و کل چهره را ساده‌تر می‌سازد. عملکرد شبکه عصبی به کار گرفته شده در این روش به این ترتیب است که ارزش سطح خاکستری هر پیکسل قرار گرفته در نیمه بالایی چهره به همراه همسایگی کوچکی از آن به شبکه اعمال می‌گردد. خروجی شبکه، کلاس (چشم و یا غیر چشم) پیکسل مرکزی این همسایگی را مشخص می‌کند. بدین ترتیب نقاط کاندیدی‌های چشم نمایان می‌گردند. در این روش از متفاوت بودن بافت چشم با سایر نواحی چهره به عنوان ویژگی در شبکه عصبی استفاده گردیده است که مهمترین مزیتش سرعت بالای آن می‌باشد. در مرحله بعد از آنالیز اجزای بهم‌پیوسته برای آماده‌سازی نواحی کاندیدی باقی‌مانده و انتخاب دقیق محل چشمها استفاده می‌گردد. ویژگی‌های ذهنی^۱ که برای فی‌لتر کردن نواحی باقی‌مانده استفاده می‌شوند عبارتند از: مساحت ناحیه^۲، وسعت افقی و عمودی هر ناحیه (که این ویژگی‌ها با حد آستانه‌ای مقایسه می‌گردند) و تراز هندسی اجزای چشم^۳ که در آن زاویه چرخش دو کاندیدی همسایه بررسی می‌گردد. در آخرین مرحله با استفاده از الگوریتم جابجایی میانگین مود غالب توزیع رنگ در ناحیه مردمکهای انتخاب شده در هر فریم از تصویر تشخیص داده می‌شود و بدین ترتیب محل مردمک ردیابی می‌گردد. این الگوریتم در فرایندی تکراری مرکز همسایگی ناحیه در حال جستجو را به بالاترین میانگین محلی هدایت می‌کند تا در نهایت همسایگی مود غالب در آن توزیع شناخته شود. این عمل تا زمانی که اختلاف بین دو بردار میانگین در دو تکرار پیاپی از حد معینی کمتر نشود ادامه می‌یابد.

روشی که توسط Fris و همکارانش پیشنهاد شد [۳۵] به تشخیص و ردیابی ویژگی‌های چهره در یک دنباله ویدیویی می‌پردازد. در این تحقیق از مدل آماری رنگ - پوست برای جداسازی نواحی کاندیدی‌های چهره در تصویر استفاده می‌گردد. با استفاده از یک آشکارساز چشم بر پایه تطبیق الگو، صحت نواحی انتخاب شده به عنوان چهره بررسی می‌گردد. پس از تشخیص چهره، مردمکها، حفره‌های بینی و کنج لبها به عنوان ویژگی استخراج شده و سپس این ویژگی‌ها در یک پروسه بلاد رنگ ردیابی می‌گردند. در این روش نیز مدل رنگ - پوست و تهیه تصویر باینری مشابه روش Kim و همکارش [۳۴] می‌باشد. به منظور حذف نقاط نویز مانند در تصویر باینری از عملگر مورفولوژی بستن^۴ و سپس الگوریتم median Filter استفاده می‌گردد. در مرحله آخر نواحی کوچک که کاندیدی‌های چهره نیستند حذف خواهند شد. نواحی باقی‌مانده ممکن است چهره و یا قسمت دیگری از بدن باشند. برای انتخاب دقیق ناحیه چهره از الگوریتم مبتنی بر تطبیق الگو (الگوی چشم) استفاده می‌گردد. از نقاط قوت این روش این است که اندازه (ابعاد) الگوی چشم که ناحیه مستطیلی است، متناسب با ابعاد ناحیه پوست مورد پردازش تعریف

¹ Heuristic

² Fill factor

³ Alignment geometric

⁴ Closing

می‌کند. بنابراین الگوریتم تطبیق الگو فقط یکبار اجرا می‌شود. بعد از شناسایی چهره، سیستم به جستجوی مردمکها، حفره‌های بینی و کنجهای لب می‌پردازد. سپس با جستجوی دو نقطه تاریک که ویژگی‌های هندسی مشابه دو مردمک دارند، مردمکها تشخیص داده می‌شوند. برای تعیین کنج لبها ابتدا محدوده تقریبی لبها از روی محل چهره و مردمکها تقریب زده می‌شود، لبه‌های افقی این ناحیه توسط عملگر sobel استخراج می‌شوند و سپس مرزهای افقی لبها با محاسبه نگاشت انتگرال عمودی شناسایی می‌گردند. مرزهای چپ و راست لبها با مقایسه این نگاشت و یک حد آستانه مشخص می‌گردند. حفره‌های بینی نیز مشابه مردمک تعیین می‌گردند. در نهایت ردیابی مردمک و حفره بینی در تصاویر وی‌دئویی، با پیدا کردن تاریکترین نقطه در اطراف مکان مردمک در فریم قبلی صورت می‌گیرد.

روشی که توسط Zhu و همکارش [۳۶] ارائه گردید به تشخیص و ردیابی چشمها تحت شرایط نوری متفاوت و چرخش سر در زوایای مختلف می‌پردازد. در این مقاله با ترکیب پیشرفتهای اخیر در شناسایی و ردیابی شیء، روشی جامع برای شناسایی و ردیابی در یک سیستم تصویر برداری مادون قرمز ارائه می‌گردد. این روش به طور قابل ملاحظه‌ای نسبت به شرایط متغیر روشنایی و چرخشهای متعدد چهره پایدار است. بعلاوه این روش عملکرد مطلوبی برای تصاویر نویزی و افرادی نکی دارد و قادر است به طور همزمان چشمهای چندین فرد با فواصل و حالات مختلف چهره را ردیابی کند. شناسایی محل چشمها با بررسی همزمان اثر روشنایی مردمک به دلیل تابش نور مادون قرمز و همچنین الگوی ظاهر چشم^۲ از طریق روش ماشین بردار پشتیبان^۳ صورت می‌گیرد.

در اولین مرحله از فرایند تشخیص چشمها برای جلوگیری از تأثیر منفی سالی منابع نوری روی مردمکها از الگوریتم پیش پردازشی استفاده می‌گردد. روشی تفاضلی که تداخلات پس زمینه را با تقریب تصاویر مربوط به مردمک تیره و مردمک روشن از بین می‌برد برای این کار استفاده می‌گردد. تصویر حاصل از این تفاضل عاری از جزئیات مزاحم پس زمینه است. سپس تصویر حاصل از این تفاضل بر اساس هیستوگرام روشنایی تصویر، آستانه‌گذاری می‌گردد و آنالیز اجزای بهم پیوسته بر روی این تصویر باینری صورت می‌گیرد تا نواحی کاندیدی مردمک (blob) شناسایی گردند. در نهایت با توجه به اینکه روشهایی از قبیل آستانه‌گذاری و تطبیق بیضی بدلیل قابل مقایسه نبودن روشنایی مردمک با سایر نقاط نوین مانند و پایدار نبودن شکل هندسی در اثر چرخش با تغییر فاصله سر نسبت به دوربین از پایداری لازم برخوردار نیستند از ماشین بردار پشتیبان (SVM) استفاده می‌گردد. با بررسی ناحیه اطراف مردمک مشاهده شده که در این نواحی توزیع روشنایی نسبتاً ثابتی وجود دارد. که نسبت به سالی نواحی چهره کاملاً متفاوت است و می‌تواند برای جداسازی چشم از غیر چشم به کار رود (شکل ۲-۳).

¹ Projection

² Eye appearance pattern

³ Support Vector Machine



شکل ۲-۳: تعیین نقاط کاندید مردمک در تصویر باینری

برای این منظور ابتدا مکانهای نواحی باقی مانده باینری، به تصویر مربوط به مردمک تیره^۱ نگاشته می شود. سپس با اعمال به یک کلاسه بند SVM نواحی متناظر با چشمها جدا می گردد.

مرحله ردیابی چشم از دو جزء اصلی تشکیل شده است. جزء اول ردیاب مرسوم با استفاده از فیلتر کالمن بر پایه مردمک روشن است. نتیجه این ردیاب، توسط کلاسه بند مبتنی بر SVM به منظور بازیابی چشمهای شناسایی شده تقویت می گردد. در مواردی که ردیاب کالمن به دلیل شدت روشنایی ضعیف در مردمک و یا عدم وجود مردمک روشن با خطا مواجه می شود از روش ردیابی مبتنی بر جابجایی میانگین^۲ برای ادامه کار استفاده می گردد. در حقیقت استفاده از این دو ردیاب ضعفهای یکدیگر را می پوشانند و پایداری سیستم را بالا می برند.

در روش ارائه شده توسط Jafari Moghadam Fard و همکارانش [۳۷]، محل مردمک بر اساس ترکیبی از الگوریتم فاصله با لایه بندی و اسکلت بندی تعیین می شود. در روش ارائه شده فرض بر این است که وضعیت سر ثابت است. سپس تصویر سطح خاکستری گرفته شده توسط دوربین آستانه گذاری شده تا تصویر باینری حاصل گردد. به منظور افزایش سرعت از آستانه عمومی ساده استفاده می گردد. تصویر خروجی شامل مردمک، مژه ها و برخی از قسمت های ابرو می باشد. مرحله بعد جداسازی مردمک از سایر بخش های تصویر است. در این مرحله گردی چشم با استفاده از برخی اطلاعات اولیه نظری فاصله بین چشم و دوربین و ابعاد چشمها قابل محاسبه خواهد بود. با اعمال الگوریتم لایه بندی^۳ و نازک سازی ناحیه مشکلی تصویر دو سطحی به محل مردمک می رسیم. این کار به حذف بخش های باریک در تصویر که به اشیاء کوچک مربوط می شوند کمک شایانی می کند. بعد از تشخیص پس پس مرکزی مربوط به مردمک، می بایست با محاسبه فاصله این نقطه تا لبه های اطراف، میزان خطا در تشخیص این نقطه را نسبت به محل واقعی مرکز مردمک بیابیم. این کار با اعمال الگوریتم فاصله به صورت معکوس^۴ با شروع از این مرکز و ختم به لبه ها صورت می گیرد. اگر اختلاف بین ارزش بی شینه و کمی نه در لبه ها برای این تصویر مساوی همین اختلاف در تصویر اولیه کالیبراسیون باشد، مرکز مردمک درست تشخیص داده شده در غیر این صورت می بایست این نقطه آنقدر شیفت داده شود تا این تساوی برقرار شود. (تصویر کالیبراسیون تصویر اولیه ای است که مرکز مردمک در آن کاملاً

¹ Dark pupil

² Mean Shift

³ Layering algorithm

⁴ Distance reverse algorithm

مشخص است). برای هر فریم جدید نیز این روش تکرار می‌گردد و محل مردمک شناسایی می‌شود.

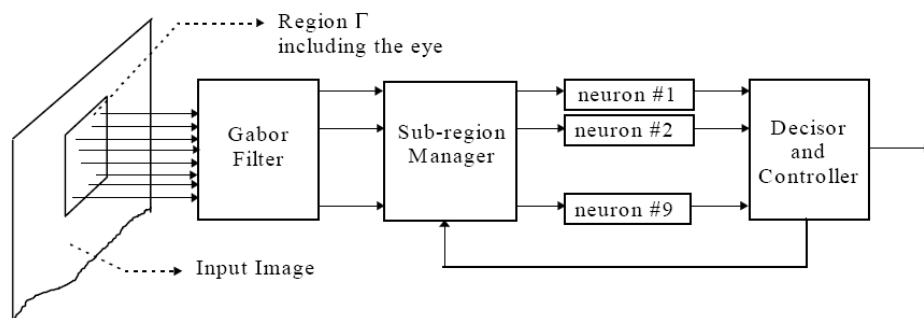
Marcone و همکارانش [۳۸] روشی برای ردیابی چشم در دنباله ویدیویی ارائه دادند که از شبکه عصبی رقابتی استفاده می‌کند. در روش آنها الگوریتمی برای ردیابی چشمها با استفاده از فیلترهای گابور محلی در تصاویر چهره ارائه می‌گردد. الگوریتم پیشنهادی توسط یک شبکه عصبی رقابتی که چشمها را به طور قابل قبولی ردیابی می‌کند پیاده‌سازی می‌گردد. ردیابی یک هدف مثل چشم در تصاویر ویدئویی با پیروی از یکی از ویژگی‌های برجسته^۱ که به طور کارآمدی با اعمال فیلترهای گابور محلی تعیین می‌شوند، امکان‌پذیر است. اعمال این الگوریتم برای تشخیص ورودی‌های چشم، قدرت عملکرد بالا و پیچیدگی محاسباتی پایینی را در پی داشت. به منظور رسی‌دن به راه بهینه برای مشخص شدن محل دقیق چشم در ناحیه مورد نظر فیلتر گابور با در نظر گرفتن دقت و پیچیدگی محاسباتی بایستی تا حد امکان بهینه گردد. برای این منظور از فیلترهای گابور متعلق به موجک با تابع پیچیده‌تری (رابطه ۲-۲۱) استفاده می‌گردد:

$$G(k, x) = e^{ik*x} e^{-(k*k*x*x)/2\delta^2} \quad 21-2$$

x بردار مربوط به یک نقطه در ناحیه فیلتر با در نظر گرفتن مرکز آن، K طول موج فاصله‌ای^۲ از اجزای $\omega \cdot \cos(\varphi)$ و $\omega \cdot \sin(\varphi)$ که فرکانس ω و چرخش فیلتر است. δ عرض پنجره گوسی را تعیین می‌کند. در این مقاله فقط بخش موهومی رابطه بالا استفاده می‌گردد چرا که به شکل موثری در استخراج لبه‌ها نقش دارد. با استفاده از چندین فیلتر با فرکانسها و چرخشهای مختلف می‌توان مجموعه کاملی از اطلاعات محلی که برای پردازش تصاویر چهره مورد نیاز است را استخراج کرد. پس از اعمال فیلتر بهینه بروی تصویر دامنه، پاسخ به صورت یک تصویر سطح خاکستری قابل نمایش است. نقطه سفید در تصویر بی‌انگرف بی‌شتترین دامنه در پاسخ بوده که محل مردمک را نشان می‌دهد. این خاصیت اساس الگوریتم ردیابی مردمک ارائه شده را می‌سازد. این شبکه (شکل ۲-۴) از ۹ نورون تشکیل شده و به طریقی زیر عمل می‌کند. نورونها پردازش را با ۹ زیر ناحیه Γ که چشم را نیز شامل می‌شود آغاز می‌کنند. در مرحله اول زیر ناحیه Γ به همراه ۸ زیر ناحیه دی‌گر از فیلتر عبور داده می‌شوند، نورون مرتبط با زیر ناحیه‌ای که ماکزیمم دامنه را در پاسخ ایجاد کند انتخاب می‌گردد. اگر این ناحیه انتخاب شده همان ناحیه مرکزی باشد این مرحله به پایان می‌رسد در غیر این صورت کنترل کننده شبکه، ناحیه انتخاب شده را به مکان مرکزی شیفت می‌دهد. دوباره این فرایند اعمال شده و نورون موفق، انتخاب می‌گردد. الگوریتم آنقدر ادامه می‌یابد که شیفتی لازم نباشد. این فرایند بر روی همه فریم‌های تصاویر اعمال می‌گردد تا در هر فریم محل مردمک تشخیص داده شود.

¹ Salient features

² Spatial



شکل ۲-۴: شبکه عصبی مورد استفاده در روش Marcone و همکارانش [۳۸]

در روش ارائه شده توسط Akhbarian Azar [۳۹]، سیستم نوروفازی^۱ با استفاده از تبدیل موجک نامتغیر با شیفت^۲ معرفی می‌شود. روش ارائه شده در این مقاله با استفاده از خوشه بندی فازی^۳ و شبکه عصبی دینامیک به استخراج اطلاعات پلک زدن می پردازد. برای محاسبه زمان تأخیر بین پلک زدن از تبدیلی موجک نامتغیر به شیفت استفاده می‌گردد. روش ارائه شده شامل سه بخش اساسی است:

۱. مرحله آموزش به منظور یادگیری شبکه برای تعداد کلاسهای پلک زدن و تهیه یک سیگنال الگو برای هر کلاس

۲. مرحله کلاسه بندی برای مشخص کردن هر پلک زدن بر اساس الگوهای ایجاد شده در مرحله قبل

۳. محاسبه تأخیر زمانی بین پلکها با استفاده از تبدیلی موجک.

مرحله یادگیری قسمت کوچکی از خوشه بندی فازی و نزدیکی همسایگی را برای استخراج الگوها استفاده می‌کند. Fuzzy C-mena و شبکه های عصبی برای کلاسه بندی پلک زدنها استفاده می‌گردد.

Tian و همکارانش [۴۰]، روشی برای تشخیص پلک زدن ارائه دادند که با استفاده از الگوی انعطاف پذیری^۴ به توصیف ویژگیهای چشم می پردازد. در الگوریتم پیشنهادی توسط آنها، دو حالت باز و بسته برای پلکها بررسی می‌شود. پس از تشخیص محل چشم در فریم اول، عنبیه با استفاده از اطلاعات شدت روشنایی و لبه شناسایی شده و ردیابی می‌گردد. در صورت شناسایی عنبیه چشم، باز تلقی شده و در غیر این صورت چشم، بسته در نظر گرفته شده و با یک خط راست مدل می‌شود. در تصویری مربوط به فریم بعدی پارامترهای پلک با استفاده از یک روش مبتنی بر ویژگی که توسط Lucas و همکارش [۴۱] ارائه شد استخراج می‌گردد. در این روش برای آشکارسازی عنبیه، ابتدا مکانی که ماسک مربوط به عنبیه (شکل ۲-۵) قویترین پاسخ را در تصویر لبه ایجاد می‌کند شناسایی می‌گردد. سپس شدت روشنایی بخشی که بالاترین پاسخ را در مرحله قبل بدست آورده است و شدت روشنایی بخش پایینی عنبیه بازبینی می‌گردند. با این حال در بسیاری از موارد نظیر روشنایی محیط، نویز و

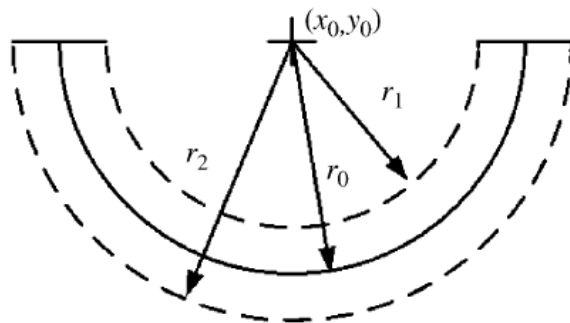
¹ Neuro-Fuzzy

² Shift invariant Wavelet

³ Fuzzy clustering

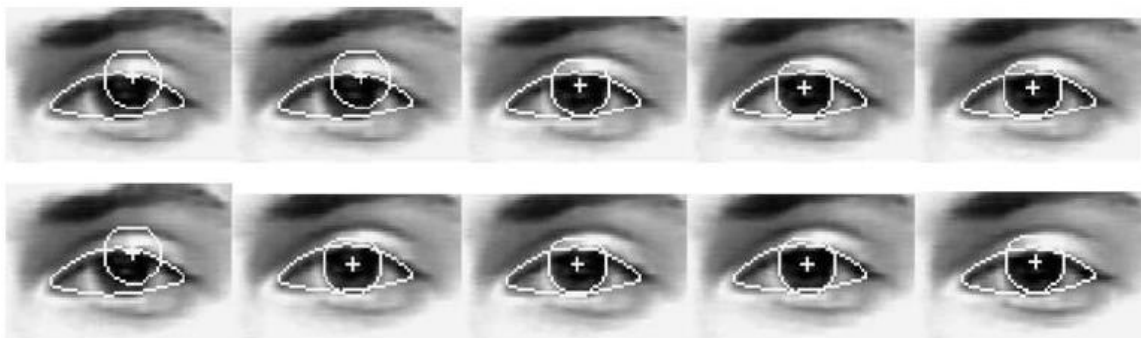
⁴ Deformable Template

پیچیدگی ناحیه چشم حتی در صورت باز بودن کامل چشم، مکان با بالاترین پاسخ همواره همان محل عنبیه نخواهد بود. نتیجتاً حالت چشم نیز اشتباه تشخیصی ص داده خواهد شد.

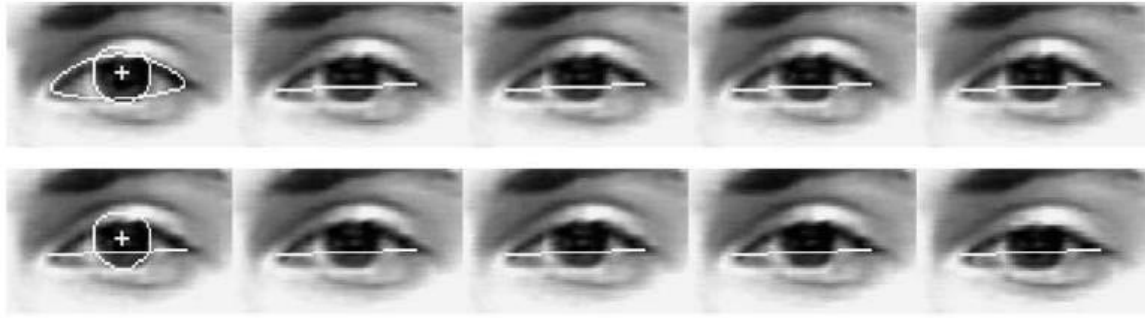


شکل ۲-۵: نیم دایره استفاده شده به عنوان ماسک: (x_0, y_0) مرکز عنبیه، r_0 شعاع عنبیه، r_1 حداقل شعاع و r_2 حداکثر شعاع ماسک

همچنین در مواردی که عنبیه دو بار شناسایی می‌گردد، ردیابی پلکها با خطا همراه است. در مدل تعریف شده برای چشم، بسته بودن چشم معادل با یک خط راست است. بنابراین در این حالت الگوهای نواحی مرکزی پلک در نیمه بالایی و پایینی با هم برابرند. از این رو هنگامی که عنبیه اشتبانه دو بار پیچیدگی شناسایی می‌گردد و یا فرد پلک می‌زند نقاط مرکزی بالا و پایینی پلک با الگوی مشابه ردیابی می‌گردند. Tan و همکارش [۴۲]، روش جدیدی برای غلبه بر دو مشکل بیان شده ارائه کردند. در روش پیشنهاد شده توسط آنها ابتدا با تغییر معیار شناسایی عنبیه مشکل تشخیصی ص عنبیه به مسأله بهینه سازی محدود مبدل می‌گردد. در فرایند بهینه سازی شدت روشنایی و اطلاعات لبه به طور همزمان و نه به ترتیب استفاده می‌گردند. علاوه بر آن، الگوهای اطراف نقاط مرکزی پلکها قبل از بسته شدن، برای پیش‌بینی الگوهای ایجاد شده هنگام باز شدن مجدد چشم مورد استفاده قرار می‌گیرد. و به این ترتیب مشکل دومی که در رابطه با روش قبل مطرح شده از بین می‌رود. روش پیشنهادی توسط Tan و همکارش [۴۲] بر روی بی‌شتر از ۵۳ تصویر متوالی افرادی که پلک می‌زنند آزمایش شد. نتایج تجربی و مقایسه با روش Tian و همکارانش [۴۰] نشان داد که روش پیشنهادی با دقت و پایداری بی‌شتری چشمها را ردیابی می‌کند. نتایج بدست آمده توسط دو روش در شکل‌های ۲-۶ و ۷-۲ نشان داده شده‌اند.



شکل ۲-۶: نتایج تشخیص و ردیابی عنبیه در روش پیشنهادی Tan و همکارش [۴۲]



شکل ۲-۷: نتایج تشخیص و ردیابی عنبیه در روش Tian و همکارانش [۴۰]

در روش ارائه شده توسط Tian و همکارانش [۴۰]، ردیابی مرزهای چشم با استفاده از دو نقطه کلی‌دی در الگوی چشم یعنی نقاط میانی پلک بالا و پایین صورت می‌گیرد. الگوی چشم‌ها نیز توسط این دو نقطه و گوشه‌های چشم تعیین می‌گردد.

همانطور که قبلاً اشاره شد نتایج این روش در زمان بسته شدن پلک‌ها، نتایج خطایی خواهند بود. برای مقابله با این مشکل فرایند ردیابی پلک‌ها در روش ارائه شده توسط Tan و همکارش [۴۲] بصورت زیر بهبود پیدا کرد.

آنها با این فرض که شدت روشنایی در فرایند پلک زدن ثابت می‌ماند، از الگوی موجود برای چشم باز قبل از پلک زدن، به منظور پیش‌بینی الگوی پلک‌ها بعد از پلک زدن (زمانیکه چشم دوباره باز می‌شود) استفاده نمودند. و برای ردیابی مقاومتر پلک‌ها از مدل AR برای پیش‌بینی نقطه مرکزی پلک استفاده کردند. رابطه ۲-۲۲ الگوی پلک‌ها را پیش‌بینی می‌کند.

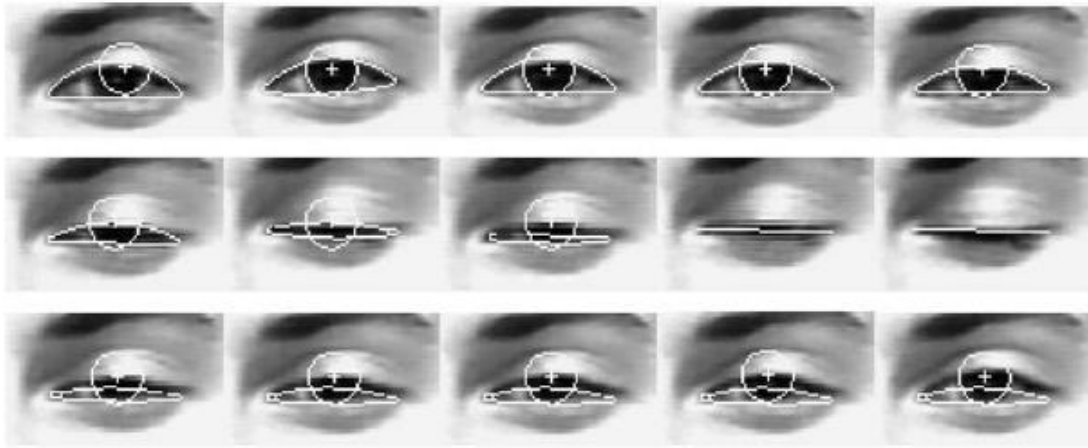
$$p_n = \sum_{i=1}^m w_i p_{n-i} \quad 22-2$$

در این رابطه p_n الگوی نقطه مرکزی پلک در فریم n ام در زمان باز بودن چشم است، m تعداد فریم‌هایی است که برای پیش‌بینی استفاده می‌گردد. و w_i نیز ضریب پیش‌گویی است.

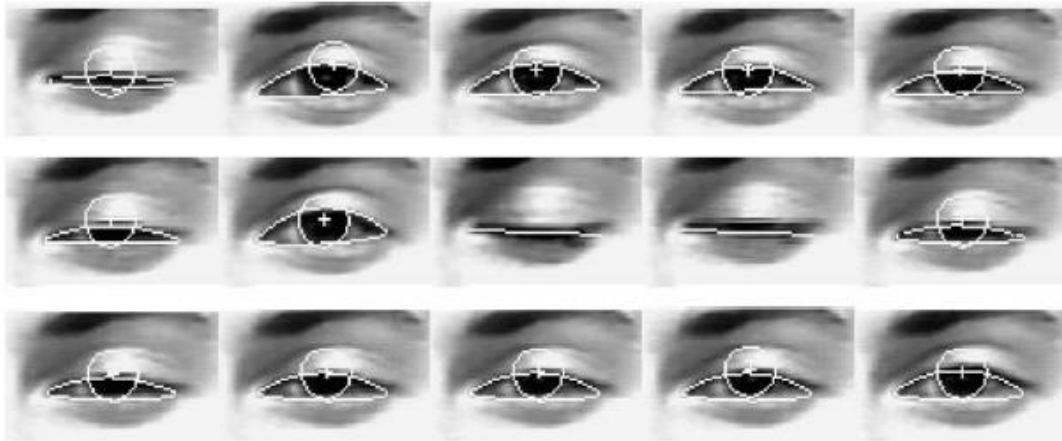
در الگوریتم ارائه شده توسط Tan و همکارش [۴۲]، مرتبه مدل ۲ قرار داده شد، بنابراین فرایند پیش‌گویی به صورت زیر قابل توصیف است:

$$p_n = w p_{n-1} + (1-w) p_{n-2} \quad 23-2$$

شکل‌های ۲-۸ و ۲-۹، نتایج ردیابی پلک‌ها را در دو الگوریتم Tian و همکارانش [۴۰] و Tan و همکارش [۴۲] نشان می‌دهد.



شکل ۲-۸: نتایج ردیابی پلک‌ها در روش Tian و همکارش [۴۰]



شکل ۲-۹: نتایج ردیابی پلک‌ها در روش Tan و همکارش [۴۲]

فصل ۳- الگوریتم پیشنهادی برای تشخیص چشم

در این فصل به تشریح عملکرد الگوریتم پیشنهادی برای تشخیص چشم می‌پردازیم. مشکلاتی که بر سر راه الگوریتمهای تشخیص چشم وجود دارد، دستیابی به پاسخ مناسب را در چنین الگوریتمهایی با دشواری روبرو می‌سازد. هدف اصلی این تحقیق ارائه الگوریتمی جدید برای تشخیص و ردیابی چشم می‌باشد. مقاوم بودن روش نسبت به تغییر نورپردازی، حضور عوامل خارجی و چرخش سر مورد توجه بوده‌است. شکل ۳-۱ مراحل مختلف الگوریتم پیشنهادی را بصورت شماتیک نشان می‌دهد.

همانطور که در شکل ۳-۱ نشان داده شده‌است، الگوریتم پیشنهادی از دو مرحله اصلی تشکیل شده‌است:

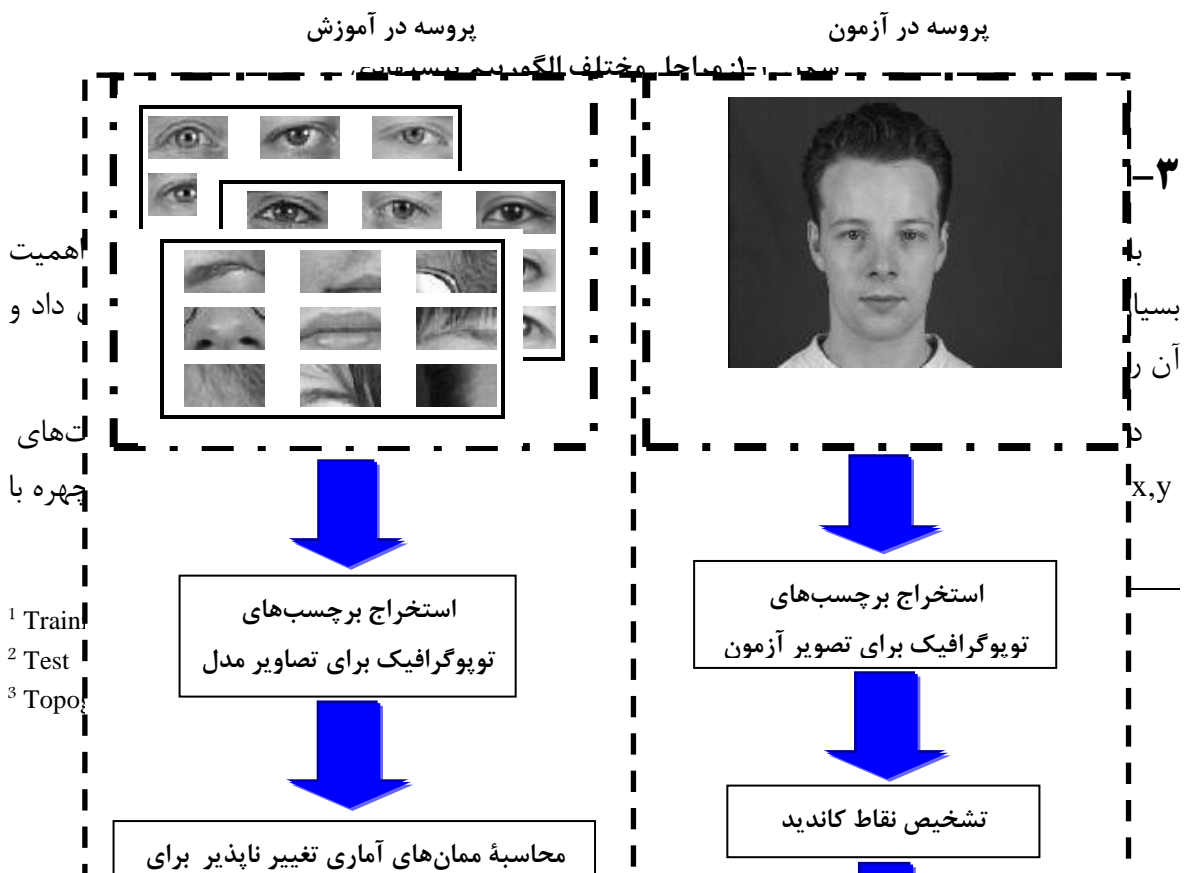
۱. مرحله آموزش^۱

۲. مرحله آزمایش^۲

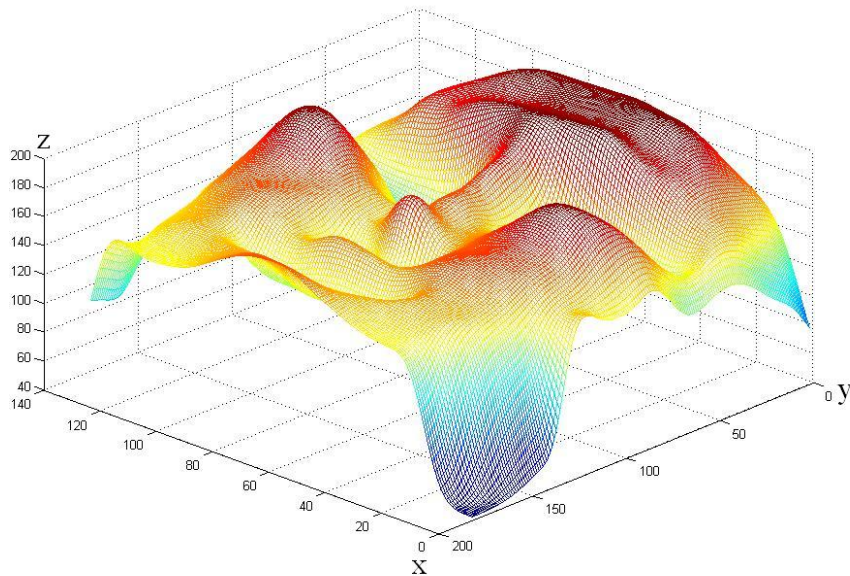
در مرحله آموزش که بصورت off line اجرا می‌شود، مدل‌های آماری سه کلاس چشم راست، چشم چپ و غیر چشم فراهم می‌گردد. تعیین مکان چشم‌ها در مرحله آزمایش با توجه به نتایج بدست آمده در مرحله اول انجام می‌شود. داده‌های ورودی به هر دو مرحله تصاویر سطح خاکستری هستند.

در مرحله آموزش به کمک تصاویر نمونه بریده شده از چشم راست، چشم چپ و غیر چشم مدل آماری این سه کلاس تعیین می‌گردد. این مرحله قبل از تشخیص مکان چشم‌ها در یک تصویر آزمون صورت می‌گیرد. لذا اجرای این مرحله بصورت off line می‌باشد.

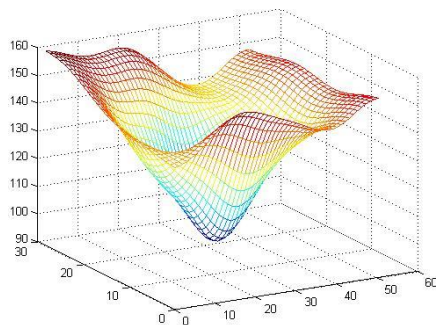
در مرحله آزمون به کمک مدل‌های آماری بدست آمده در مرحله آموزش، تصویر آزمون را آنالیز نموده تا به مکان چشم‌ها دست یابیم. تصویر آزمون، تصویر سطح خاکستری یک فرد می‌باشد که دو چشم در آن حضور دارند. اجرای این مرحله در کاربردهای زمان واقعی بایستی سریع باشد.



توجه به مشخصات توپوگرافیکی (پستی و بلندی‌ها) در رویه متناظر با آن، قابل تشخیص می‌باشند. تصویر سطح خاکستری یک چهره و رویه متناظر با آن در شکل ۲-۳ نمایش داده شده‌است.



شکل ۲-۳: نمایش تصویر چهره بصورت یک رویه سه بعدی



شکل ۳-۳: رویه متناظر با چشم راست

همانطور که در شکل ۳-۳ مشخص است، وجود مینیم محلی در مردمک و دامنه‌هایی که در اطراف آن قرار گرفته‌اند، چشم را از دیگر اجزاء چهره متمایز می‌کند.

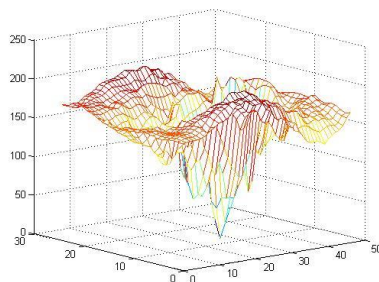
برای استخراج ویژگی‌های توپوگرافیکی از یک تصویر، به هر پیکسل از تصویر یکی از دوازده برجسب توپوگرافیکی را نسبت می‌دهیم. این برجسب‌ها وضعیت رویه را در نقطه مورد نظر مشخص می‌کنند، نظیر قلّه، چاله، سراشیبی و

دلایل قابل توجهی برای استفاده از تصویر نگاشت یافته به حوزه توپوگرافیک به منظور تعیین ویژگی‌های چشم وجود دارد، مهمترین آن این است که تصویر نگاشت یافته در مقایسه با تصویر سطح خاکستری نسبت به شرایط نورپردازی مقاوم‌تر است. در تصویر سطح خاکستری سطح روشنایی پیکسل‌ها در ۲۵۶ سطح (۰-۲۵۵) کوانتیزه شده‌است، این مقادیر در اثر تغییر شرایط تصویر برداری و نورپردازی تغییر می‌کنند در حالیکه در تصویر برچسب خورده توپوگرافیکی ارزش پیکسل‌ها به ۱۲ برچسب که توصیف‌گر سطح رویه می‌باشند، خلاصه می‌شود. این مقادیر بیانگر خاصیت رویه در محل پیکسل می‌باشند و مقاومت بیشتری در مقابل تغییر شرایط تصویر برداری و نورپردازی دارند. همچنین وجود مینیمم محلی در مردمک که در تصویر برچسب خورده ظاهر می‌شود ویژگی قوی و مناسبی برای تعیین نقاط کاندیدای مردمک و بررسی ناحیه اطراف آن برای وجود چشم می‌باشد.

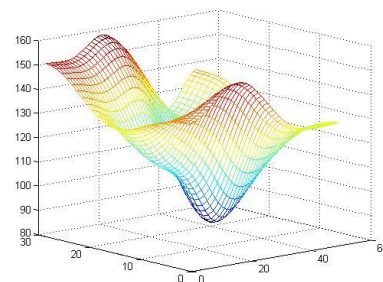
۳-۱-۱- حذف نویز از تصویر ورودی

استخراج برچسب‌های توپوگرافیک تصویر نیازمند محاسباتی است که از مشتقات مرتبه اول و دوم در هر نقطه از رویه تصویر استفاده می‌کند. در صورتیکه تصویر نویزی باشد، مقادیر محاسبه شده برای مشتقات بنحو قابل توجهی تحت تأثیر نویز قرار گرفته و منجر به تولید نتایج متفاوت در ساختار توپوگرافیک تصویر می‌شود.

روشی که برای مقابله با اثرات ناخواسته نویز وارد شده در تصاویر بکار گرفته می‌شود، هموارسازی^۱ رویه قبل از محاسبه مشتق می‌باشد. پروسه هموارسازی با استفاده از یک فیلتر گوسی دوبعدی^۲ امکان‌پذیر می‌باشد. شکل ۳-۴ رویه یک تصویر نمونه را قبل و بعد از هموار سازی نشان می‌دهد.



(الف)



(ب)

شکل ۳-۴: الف) تصویر اصلی چشم و رویه مربوط به آن (ب) تصویر چشم پس از اعمال فیلتر گوسی و رویه مربوط به آن

¹ Smoothing

² Gaussian filter

نتایج بدست آمده نشان می‌دهد که انتخاب پارامتر انحراف معیار فیلتر گوسی (δ) از اهمیت زیادی برخوردار است. بزرگ بودن انحراف معیار باعث هموارسازی شدید و از دست رفتن مشخصات توپوگرافی کی‌چهره می‌شود در حالی که انتخاب مقدار کوچک برای δ اثرات نویز و تغییرات در شرایط تصویر برداری را حذف نمی‌کند. وجود جزئیات زیاد در تصویر عدم سازگاری بین تصاویر برچسب خورده چشم در شرایط و افراد مختلف را باعث می‌شود.

در الگوریتم پیشنهادی سعی شده است با اعمال فیلتر گوسی در حالیکه جزئیات و نویز در رویه متناظر با تصویر حذف می‌گردد، ویژگی‌های توپوگرافیکی آن حفظ شود. برای این منظور مقدار مناسبی برای انحراف معیار فیلتر گوسی (δ) تعیین می‌شود.

۳-۱-۲- استخراج برچسب‌های توپوگرافیک

همانطور که قبلاً نیز بیان شد، برای دست‌یابی به ساختار توپوگرافیک تصویر باید هریک از پیکسل‌های تصویر بر اساس حالت رویه در آن نقطه برچسب گذاری شود.

الگوریتم‌های متفاوتی برای استخراج ساختار توپوگرافیک یک تصویر دیجیتال در تارای‌خچه بی‌نالی ماشینی گزارش شده است [۴۳]. این الگوریتم‌ها به دو دسته اصلی تقسیم می‌شوند. دسته اول به تکنیک‌هایی اشاره می‌کند که بطور مستقیم از تصویر دیجیتال (گسسته) برای طبقه‌بندی پیکسل‌ها استفاده می‌کنند در حالیکه در دسته دوم ابتدا یک سطح پیوسته بطور محلی در هر نقطه از رویه با شرط ایجاد مینیمم خطا، برازش^۱ شده و سپس طبقه بندی پیکسل‌ها بر اساس رویه تخمین زده شده انجام می‌شود [۴۴].

در این تحقیق مشابه روش Wang [۴۵]، چندجمله‌ای گسسته چبی‌چف^۲ بعنوان اساس گسترش فضای برداری این توابع پیوسته انتخاب شد. مشتق جزئی مرتبه (p,q) تابع f در نقطه (x,y) ، در راستای محور x و q در راستای محور y توسط رابطه ۳-۱ محاسبه می‌شود:

$$f^{(p,q)}(x,y) = \sum_{i=-N}^N \sum_{j=-N}^N f(x-i, y-j) h(i,p) h(j,q) \quad 1-3$$

در این رابطه $f(x,y)$ تصویر سطح خاکستری ورودی است و $h(i,p)$ و $h(j,q)$ فیلترهای متمایز هموار شده به ترتیب با درجه‌های p و q از چندجمله‌ای‌های چبی‌چف هستند [۴۶]. روابط ۳-۲، ۳-۳ و ۳-۴ به ترتیب $h(i,0)$ ، $h(i,1)$ و $h(i,2)$ را نشان می‌دهند.

$$h(i,0) = -\frac{3[5i^2 - (3N^2 + 3N - 1)]}{(2N - 1)(2N + 1)(2N + 3)} \quad i = -N \dots N \quad 2-3$$

$$h(i,1) = \frac{5[7(3N^2 + 3N - 1)i^3 - 5(3N^4 + 6N^3 - 3N + 1)i]}{(N - 1)N(N + 1)(N + 2)(2N - 1)(2N + 1)(2N + 3)} \quad i = -N \dots N \quad 3-3$$

¹ Fit

² Discrete Chebyshev polynomials

$$h(i,2) = \frac{30[3i^2 - N(N+1)]}{N(N+1)(2N-1)(2N+1)(2N+3)} \quad i = -N \dots N \quad 4-3$$

با توجه به اینکه فیلتر h مستقل از تصویر ورودی می باشد، می تواند بطور مستقل برای ابعاد همسایگی (N) انتخاب شده، محاسبه گردد. مشتق جزئی مرتبه $(p+q)$ از کانالو فیلتر گسسته دوبعدی با تصویر بدست می آید (رابطه ۵-۳).

$$f^{(p,q)}(x,y) = conv(f(x,y), H(p,q)) \quad 5-3$$

در این رابطه $H(p,q)$ فیلتر مشتق گیر با درجه های p و q را نشان می دهد. p درجه مشتق در راستای محور x و q درجه مشتق در راستای محور y می باشد. فیلتر H از ضرب بردارهای h ساخته می شود و ابعاد آن نیز $N \times N$ می باشد.

روابط ۶-۳ تا ۱۰-۳ فیلترهای مورد استفاده در الگوریتم پیشنهادی را نشان می دهند.

$$H(1,0) = h^T(i,1) \times h(i,0) \quad 6-3$$

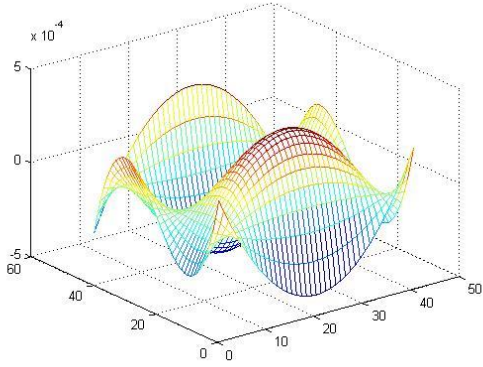
$$H(0,1) = h^T(i,0) \times h(i,1) \quad 7-3$$

$$H(1,1) = h^T(i,1) \times h(i,1) \quad 8-3$$

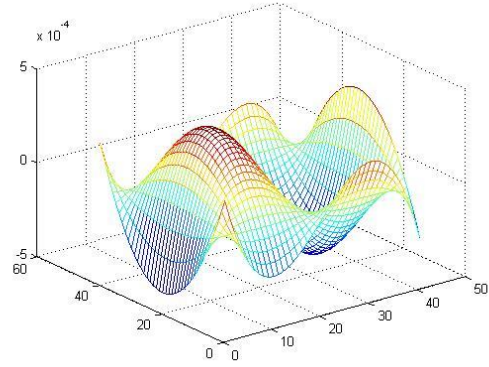
$$H(2,0) = h^T(i,2) \times h(i,0) \quad 9-3$$

$$H(0,2) = h^T(i,0) \times h(i,2) \quad 10-3$$

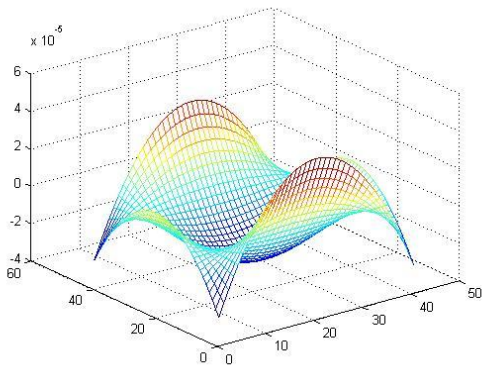
شکل ۵-۳ فیلترهای بدست آمده برای $N=20$ را نشان می دهند.



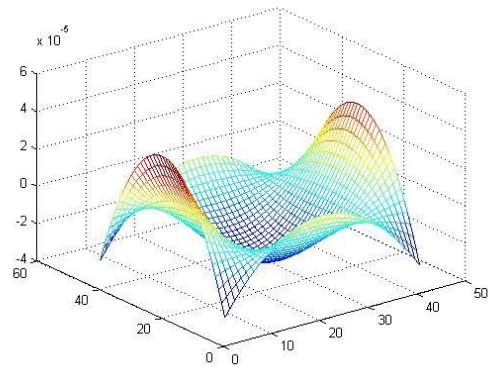
$H(1,0)$: (الف)



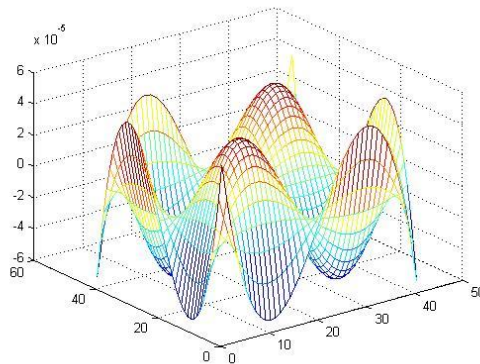
$H(0,1)$: (ب)



$H(2,0)$: (ج)



$H(0,2)$: (د)



$H(1,1)$: (ه)

شکل ۳-۵ : فیلترهای گسسته دو بعدی برای $N=20$

بردار گرادیان^۱ و اندازه آن در نقطه‌ای به مختصات (x,y) ، توسط روابط ۳-۱۱ و ۳-۱۲ بیان می‌شوند. در این روابط $\frac{\partial f(x,y)}{\partial x}$ و $\frac{\partial f(x,y)}{\partial y}$ مشتق جزئی تابع f نسبت به x و y می‌باشند.

^۱ Gradient vector

$$\vec{\nabla}f = \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y} \right) \quad 11-3$$

$$\|\vec{\nabla}f\| = \sqrt{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2} \quad 12-3$$

ماتریس Hessian نیز برای نقطه (x,y) از تصویر بصورت زیر تعریف می‌گردد:

$$H = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial x \partial y} \\ \frac{\partial^2 f(x,y)}{\partial x \partial y} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} f^{(2,0)}(x,y) & f^{(1,1)}(x,y) \\ f^{(1,1)}(x,y) & f^{(0,2)}(x,y) \end{bmatrix} \quad 13-3$$

بردارهای ویژه^۱ این ماتریس \bar{u}_1, \bar{u}_2 و مقادیر ویژه^۲ متناظر با این بردارها λ_1, λ_2 هستند و شرط $|\lambda_1| \geq |\lambda_2|$ برقرار می‌باشد.

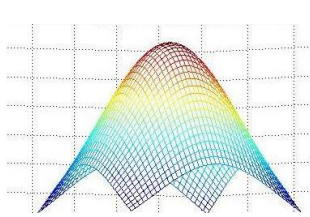
دو نکته مهم در رابطه با $\bar{u}_1, \bar{u}_2, \lambda_1, \lambda_2$ وجود دارد:

۱. H یک ماتریس متقارن است بنابراین λ_1, λ_2 حقیقی هستند و \bar{u}_1, \bar{u}_2 برهم عمودند.
۲. مشتق دوم f در جهت \bar{u} بصورت $f''_{\bar{u}} = \bar{u}^T H \bar{u}$ بیان می‌شود. بردارهای ویژه ماتریس H ، \bar{u}_1, \bar{u}_2 جهت‌هایی هستند که مشتق دوم در آنها ماکزیمم است. λ_1, λ_2 نیز بیانگر مقادیر مشتق دوم هستند.

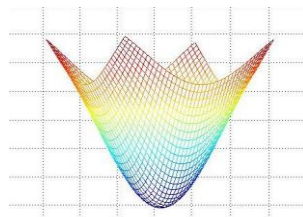
۳. $\vec{\nabla}f \cdot \bar{u}_1$ ، تصویر بردار گرادیان در جهت \bar{u}_1 و $\vec{\nabla}f \cdot \bar{u}_2$ تصویر بردار گرادیان در جهت \bar{u}_2 است.

مقادیر محاسبه شده برای $\|\vec{\nabla}f(x,y)\|, \bar{u}_1, \bar{u}_2, \lambda_1, \lambda_2$ ، نوع برجسب توپوگرافیکی که به هر پیکسل اختصاص می‌یابد را مشخص می‌کنند.

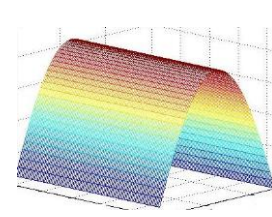
۱۲ برجسب توپوگرافیک در شکل ۳-۶ نشان داده شده‌اند.



(الف): قله



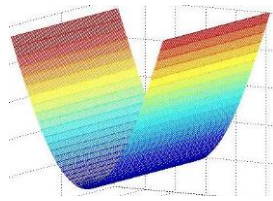
(ب): چاله



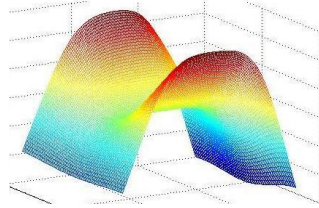
(ج): ناودانی بیرون

¹ Eigen vectors

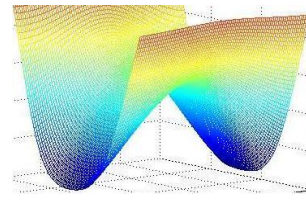
² Eigen values



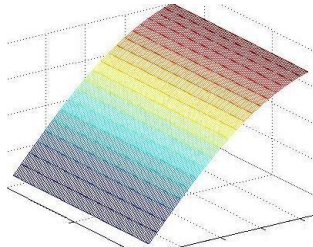
(د): ناودانی داخل



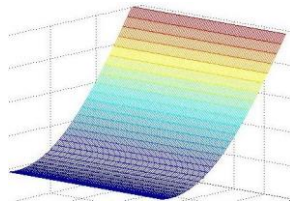
(ه): زین اسبی از خارج



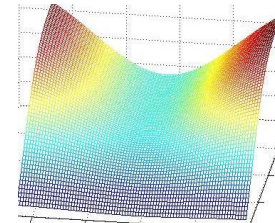
(ز): زین اسبی از داخل



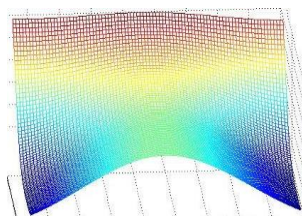
(ح): سرایشی با انحناء به داخل



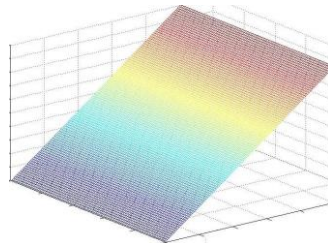
(ط): سرایشی با انحناء به خارج



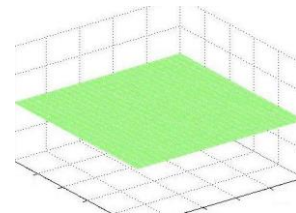
(ی): سرایشی زین اسبی با انحناء به خارج



(ک): سرایشی زین اسبی با انحناء به داخل



(ل): سرایشی بدون انحناء



(م): سطح تراز

شکل ۳-۶: برچسب‌های توپوگرافیک

قله^۱: اولین برچسب از برچسب‌های دوازده گانه می‌باشد و زمانی اتفاق می‌افتد که یک ماکزیمم محلی^۲ در تمام جهات وجود داشته باشد. شرایط بیان شده در رابطه ۳-۱۴ نقطه به مختصات (x,y) از تصویر را در گروه قله قرار می‌دهد [۴۴].

$$\left\| \vec{\nabla} f(x,y) \right\| = 0 \quad , \quad \lambda_1 < 0, \lambda_2 < 0 \quad 14-3$$

چاله^۳: به برچسب شماره دو اشاره دارد و زمانی رخ می‌دهد که یک مینیمم محلی^۴ در تمام جهات وجود داشته باشد. در این برچسب مشابه قله، اندازه گرادیان صفر می‌باشد اما مقادیر ویژه مثبت هستند [۴۴] (رابطه ۳-۱۵).

¹ Peak

² Local maximum

³ Pit

⁴ Local minimum

$$\left\| \vec{\nabla} f(x, y) \right\| = 0, \quad \lambda_1 > 0, \lambda_2 > 0$$

۱۵-۳

ناودانی بیرون^۱: سومین برجسب توپوگرافیکی می‌باشد و به پیکسلی الصاق می‌گردد که در آن نقطه، ماکزیمم محلی در یک جهت وجود داشته باشد. از به هم پیوستن نقاط ناودانی بیرون یک خط بوجود می‌آید که این خط ممکن است تراز^۲، همراه با شیب رو به بالا^۳، شیب رو به پایین^۴، خمیدگی رو به بالا^۵ یا خمیدگی رو به پایین^۶ باشد. نقطه^۷ (x,y) در تصویر، برجسب ناودانی بیرون می‌خورد در صورتیکه یکی از شرایط بیان شده در رابطه ۳-۱۶ در مورد آن صدق کند [۴۴].

$$\left\| \vec{\nabla} f(x, y) \right\| = 0, \quad \lambda_1 < 0, \lambda_2 = 0$$

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 < 0, \lambda_2 = * , \quad \vec{\nabla} f \cdot \vec{u}_1 = 0$$

۱۶-۳

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 = *, \lambda_2 < 0, \quad \vec{\nabla} f \cdot \vec{u}_2 = 0$$

ناودانی داخل^۷: ناودانی داخل همانند ناودانی بیرون است با این تفاوت که در ناودانی داخل یک مینیمم محلی در یک جهت اتفاق می‌افتد. رابطه ۳-۱۷ شرایط عنوان شده برای این برجسب را نشان می‌دهد [۴۴].

$$\left\| \vec{\nabla} f(x, y) \right\| = 0, \quad \lambda_1 > 0, \lambda_2 = 0$$

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 > 0, \lambda_2 = * , \quad \vec{\nabla} f \cdot \vec{u}_1 = 0$$

۱۷-۳

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 = *, \lambda_2 > 0, \quad \vec{\nabla} f \cdot \vec{u}_2 = 0$$

زین اسبی از داخل^۸ و خارج^۹: یکی از این دو برجسب زمانی به یک پیکسل اختصاص می‌یابد که در آن نقطه یک ماکزیمم محلی و یک مینیمم محلی در دو جهت عمود برهم وجود داشته باشد. بنابراین اندازه گرادیان در نقطه مورد نظر صفر می‌باشد و علامت مشتقات مرتبه دوم در آن نقطه باید مخالف هم باشند.

¹ Ridge

² Flat

³ Slope upward

⁴ Slope downward

⁵ Curve upward

⁶ Curve downward

⁷ Ravine

⁸ Ravine Saddle

⁹ Ridge Saddle

روابط ۱۸-۳ و ۱۹-۳ به ترتیب شرایط حاکم بر زین اسبی از خارج و زین اسبی از داخل را بیان می‌کنند [۴۴].

$$\left\| \vec{\nabla} f(x, y) \right\| = 0, \quad \lambda_1 < 0, \lambda_2 > 0 \quad 18-3$$

$$\left\| \vec{\nabla} f(x, y) \right\| = 0, \quad \lambda_1 > 0, \lambda_2 < 0 \quad 19-3$$

دامنه^۱: پنج برجسب بعدی در گروه دامنه‌ها قرار می‌گیرند. وجه مشترک تمام آنها مخالف صفر بودن اندازه گرادیان در نقطه مورد تحلیل می‌باشد. روابط ۲۰-۳ تا ۳۰-۳ شرایط مربوط به هریک از برجسبها در این گروه را بیان می‌کنند [۴۴].

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 < 0, \lambda_2 < 0 \quad 20-3 \quad \text{سراشیبی با انحناء به خارج}$$

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 < 0, \lambda_2 = 0$$

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 > 0, \lambda_2 > 0$$

$$21-3 \quad \text{سراشیبی با انحناء به داخل}$$

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 > 0, \lambda_2 = 0$$

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 < 0, \lambda_2 > 0$$

$$22-3 \quad \text{سراشیبی زین اسبی با انحناء به خارج}$$

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 > 0, \lambda_2 < 0$$

$$23-3 \quad \text{سراشیبی زین اسبی با انحناء به داخل}$$

$$\left\| \vec{\nabla} f(x, y) \right\| \neq 0, \quad \lambda_1 = 0, \lambda_2 = 0$$

$$24-3 \quad \text{سراشیبی بدون انحناء}$$

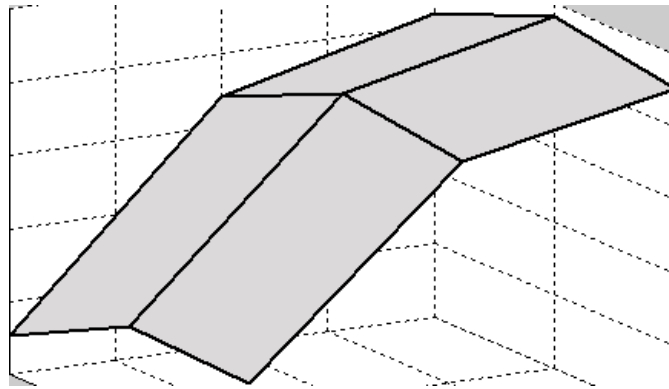
سطح تراز^۲: برجسبی است که به نقاط متعلق به رویه مسطح و تراز با سطح افق الصاق می‌گردد. رابطه ۲۵-۳ شرط مربوط به این برجسب را نشان می‌دهد [۴۴].

¹ Hillside

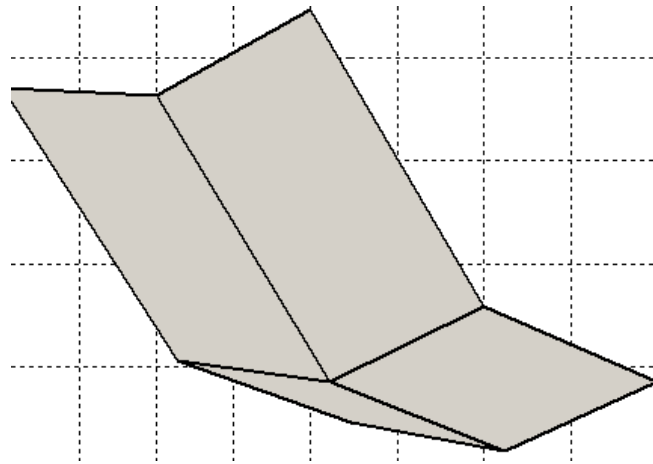
² Flat

$$\|\vec{\nabla}f(x,y)\| = 0, \quad \lambda_1 = 0, \lambda_2 = 0$$

نتایج حاصله نشان می‌دهد که بدلیل گسسته بودن رویه‌ها در هنگام اعمال شرط‌های مربوط به هر برچسب با مشکل مواجه خواهیم شد و کلاسه‌بندی نتایج صحیحی نمی‌دهد. معمولاً برچسب‌های توپوگرافیکی دقیقاً در مرکز پیکسل اتفاق نمی‌افتند (شکل ۷-۳ و شکل ۸-۳)، بنابراین اگر برچسب پیکسل واقع در نقطه مرکزی به آن پیکسل اختصاص یابد باعث ایجاد خطا در کلاسه‌بندی پیکسل‌های تصویر خواهد شد.



شکل ۷-۳: وقوع ماکزیمم محلی در مکانی بجز مرکز پیکسل



شکل ۸-۳: وقوع مینیمم محلی در مکانی بجز مرکز پیکسل

در روش ارائه شده توسط Haralik و همکارانش [۴۴] برای حل این مشکل در نقطه مورد تحلیل، به جستجوی اکستریم‌ها در جهت بردارهای ویژه پرداخته می‌شود. اما این روش بدلیل پیچیدگی محاسباتی زیاد سرعت برچسب زنی را تا زیادی کاهش می‌دهد. Wang و همکارش [۴۵] با تغییر این روش بصورت زیر مقداری از محاسبات آن را کاهش دادند.

الگوریتم مورد استفاده توسط Wang و همکارش [۴۵] برای تعیین برچسب توپوگرافیک در پیکسل P با بردارهای ویژه \vec{u}_1, \vec{u}_2 و مقادیر ویژه λ_1, λ_2 بصورت زیر می‌باشد:

- اگر اندازه گرادیان در پیکسل P صفر باشد، برچسب پیکسل با توجه به شرایط بیان شده در مورد λ_1, λ_2 بطور مستقیم تعیین می شود.

- در صورتیکه اندازه گرادیان در پیکسل P مخالف صفر باشد مراحل زیر به ترتیب انجام می شوند:
 ۱. بردارهای ویژه (\bar{u}_1, \bar{u}_2) دارای دو جهت عمود برهم می باشند که در دو جهت از چهار جهت $0^\circ, 90^\circ, +45^\circ, -45^\circ$ تقریب زده می شوند، بردارهای نشان داده شده در شکل ۹-۳(الف) معرف این جهت ها می باشند.



(ب): نمایش بردارهای ویژه در کنار جهت های تعریف شده روی تصویر

(الف): جهت های تعریف شده روی تصویر

شکل ۹-۳: تصحیح بردارهای ویژه

۲. نزدیکترین جهت های چهارگانه به بردارهای \bar{u}_1, \bar{u}_2 به عنوان تقریب این بردارها در نظر گرفته می شوند. برای مثال در شکل ۹-۳ (ب)، این زوایا جهت های $+45^\circ, -45^\circ$ می باشند (جهت های d_1, d_2). سپس در جهت های تصحیح شده دو همسایگی برای پیکسل P ، تعیین می شود: P_{i1}, P_{i2}

۳. برای پاسخ به این سؤال که آیا در پیکسل P اکسترمم اتفاق می افتد یا نه، مشتق مرتبه اول در جهت های تخمین زده شده محاسبه می گردد. حالات مختلف پاسخ به همراه برچسب انتخابی در جدول ۱-۳ بیان شده است.

(علامت * در این جدول به مفهوم هرچه باشد است)

جدول ۱-۳: تعیین برجسب‌های توپوگرافیک تصویر با استفاده از روش Wang و همکارش [۴۵]

برجسب انتخابی	λ_2	λ_1	آیا مشتق رویه در پیکسل P و در جهت d_2 تغییر علامت داده است؟	آیا مشتق رویه در پیکسل P و در جهت d_1 تغییر علامت داده است؟
<p><i>Convex hill</i></p> <p><i>Convex Saddle hill</i></p> <p><i>Concave Saddle hill</i></p> <p><i>Convex hill</i></p> <p><i>Slope hill</i></p>	- یا 0 + - + یا 0 0	- - + + 0	خیر	خیر
<p><i>Ravine</i></p> <p><i>Ridge</i></p> <p>نوع برجسب (<i>ravin</i> یا <i>ridge</i>) با مقایسه مقدار پیکسل P و میانگین مقادیر روشنایی در دو همسایگی‌اش تعیین می‌گردد.</p>	+ - 0	× × ×	خیر	بله
<p><i>Ravine</i></p> <p><i>Ridge</i></p> <p>نوع برجسب (<i>ravin</i> یا <i>ridge</i>) با مقایسه مقدار پیکسل P و میانگین مقادیر روشنایی در دو همسایگی‌اش تعیین می‌گردد.</p>	× × ×	+ - 0	بله	خیر
<p><i>Peak</i></p> <p><i>Pit</i></p> <p><i>Ridge</i></p> <p><i>Ravine</i></p> <p><i>Ridge Saddle</i> یا <i>Ravin Saddle</i></p>	- + 0 0 ±	- + - + ±	بله	بله

۳-۱-۳- تعیین ساختار توپوگرافیک نمونه‌های آموزشی

با توجه به مطالب بیان شده در ابتدای فصل، اولین مرحله در پروسه استخراج ویژگی‌ها تعیین ساختار توپوگرافیک تصویر می‌باشد. تصاویر آموزشی در سه گروه قرار می‌گیرند:

۱. چشم راست
۲. چشم چپ
۳. غیر چشم

شکل ۳-۱۰ نمونه‌ای از این تصاویر را نشان می‌دهد.



(د) تصاویر غیر چشم

(ب): تصاویر چشم چپ

(الف): تصاویر چشم راست

شکل ۳-۱۰: تصاویر مورد استفاده در آموزش سیستم

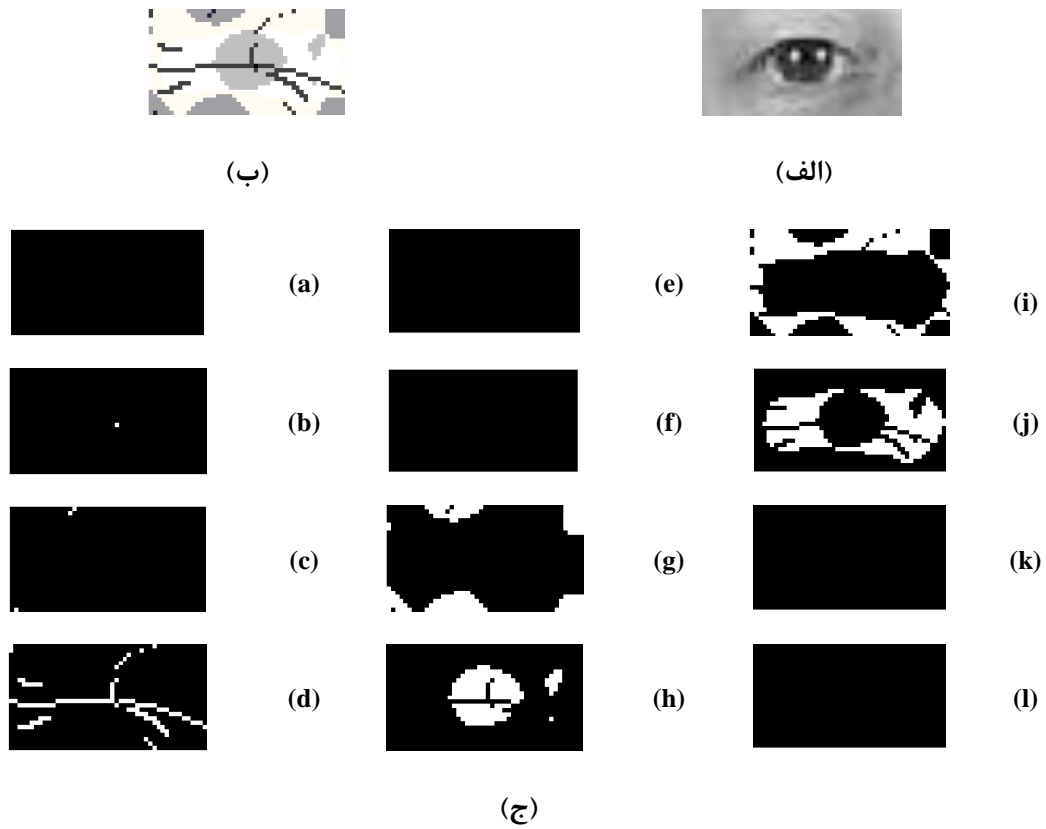
همانطور که قبلاً نیز گفته شد تصویر ناحیه چشم در رویه چهره به صورت یک مینیمم محلی در مردمک و دامنه‌هایی که آن را در بر گرفته‌اند دیده می‌شود، الگوی توصیف شده در بین اجزاء چهره منحصر بفرد می‌باشد. با نگاشت تصویر به حوزه توپوگرافیک و استخراج برچسب هر پیکسل، تصویر ناحیه‌بندی شده توپوگرافیکی حاصل می‌گردد. نتایج حاصل از کلاسه‌بندی تصویر چشم بصورت برچسب‌های توپوگرافیک، وجود برچسب‌های زیر را در ناحیه چشم نشان می‌دهد:

۱. یک مینیمم محلی در مردمک (چاله)
۲. یک دره باریک که حالت چشم را در راستای افقی چهره مشخص می‌کند (ناودانی داخل)
۳. دامنه‌ها در اطراف چاله و ناودانی داخل

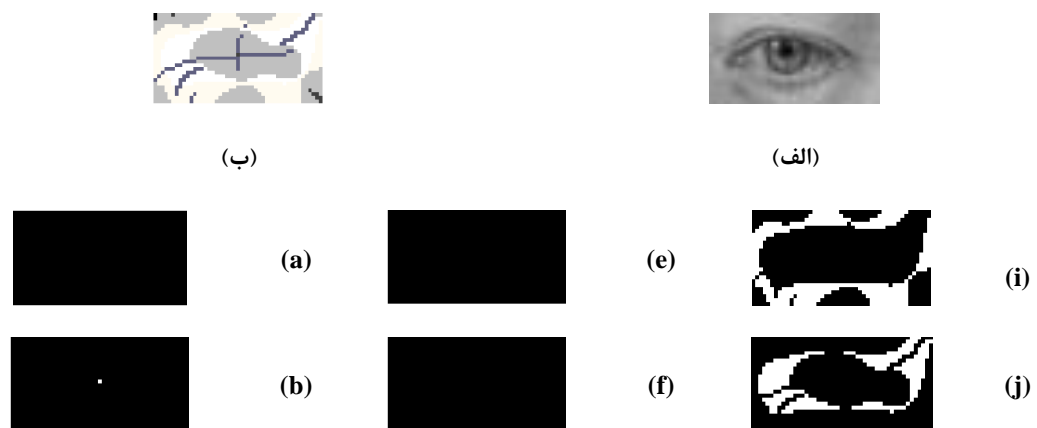
در صورتیکه تصویر را بصورت یک ماتریس دو بعدی در نظر بگیریم، تصویر داده شده برچسب‌های توپوگرافیکی را با یک ماتریس نمایش می‌دهیم که عنصر سطر x ام و ستون y ام برچسب اختصاص داده شده به پیکسل موجود در مختصات (x,y) تصویر سطح خاکستری می‌باشد. در صورتیکه این ماتریس را بصورت یک تصویر سطح خاکستری^۱ نشان دهیم، نواحی مربوط به برچسب‌های مختلف بصورت مجزا در آن مشاهده می‌شوند. نتایج برای چهار تصویر نمونه که به نمایندگی از سه کلاس چشم راست، چشم چپ

^۱ Gray Scale Image

و غیر چشم انتخاب شده‌اند، به ترتیب در شکل‌های ۱۱-۳، ۱۲-۳، ۱۳-۳ و ۱۴-۳ دیده می‌شوند. بطوریکه نواحی a تا l به ترتیب متناظر با دوازده برجسب توپوگرافیکی تعریف شده در شکل ۶-۳ می‌باشند.



شکل ۱۱-۳: الف) تصویر اولیه از چشم راست ب) تصویر نگاشت یافته به حوزه توپوگرافیک ج) تفکیک برجسب‌های ظاهر شده در تصویر نگاشت یافته





(ج)

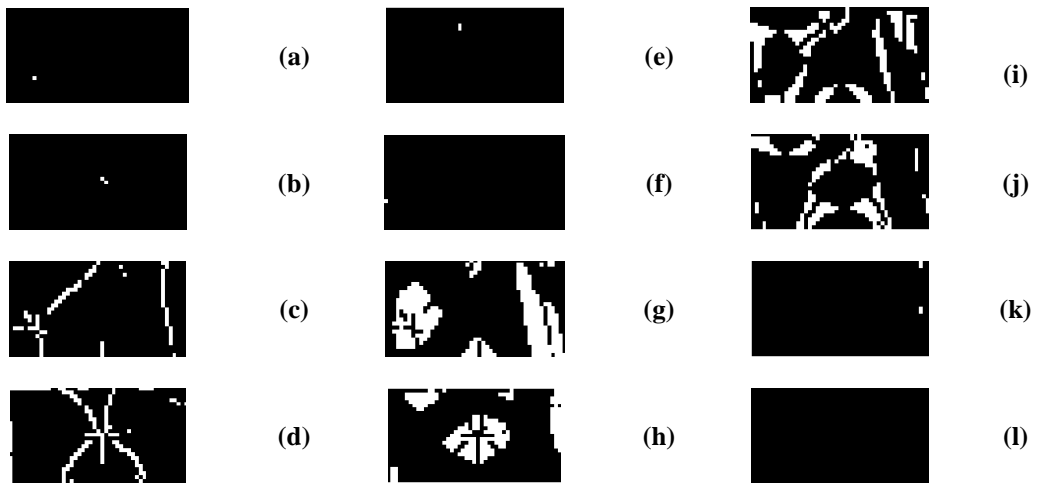
شکل ۳-۱۲: الف) تصویر اولیه از چشم چپ ب) تصویر نگاشت یافته به حوزه توپوگرافیک ج) تفکیک برجسب‌های ظاهر شده در تصویر نگاشت یافته



(ب)



(الف)



(ج)

شکل ۳-۱۳: الف) تصویر اولیه غیر چشم ب) تصویر نگاشت یافته به حوزه توپوگرافیک ج) تفکیک برجسب‌های ظاهر شده در تصویر نگاشت یافته



(ب)



(الف)



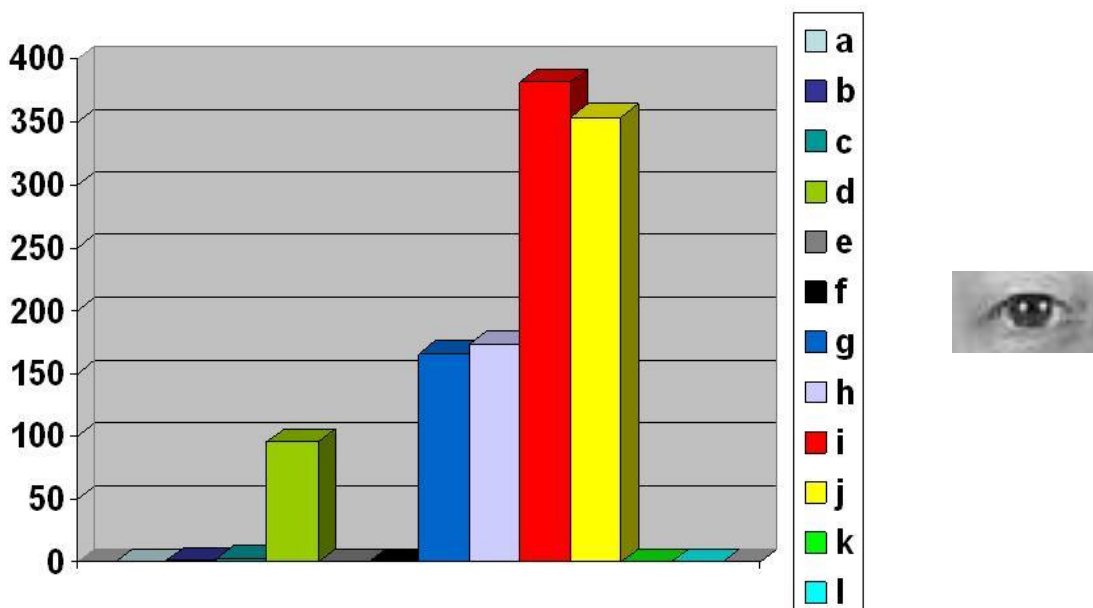


(ج)

شکل ۳-۱۴: الف) تصویر اولیه غیر چشم ب) تصویر نگاشت یافته به حوزه توپوگرافیک ج) تفکیک برجسب‌های ظاهر شده در تصویر نگاشت یافته

با توجه به نتایج حاصله از پیاده‌سازی الگوریتم بر روی پایگاه‌های داده شامل تصاویر سه کلاس مذکور مشاهده می‌گردد که برجسب‌های قلّه، زین اسبی از داخل، زین اسبی از خارج، سرایشی بدون انحناء و سطح تراز در نتایج مربوط به تصاویر چشم ظاهر نمی‌شوند. همانطوریکه انتظار می‌رود نواحی مربوط به برجسب‌های مختلف در تصاویر غیر چشم از الگوی خاصی پیروی نمی‌کنند.

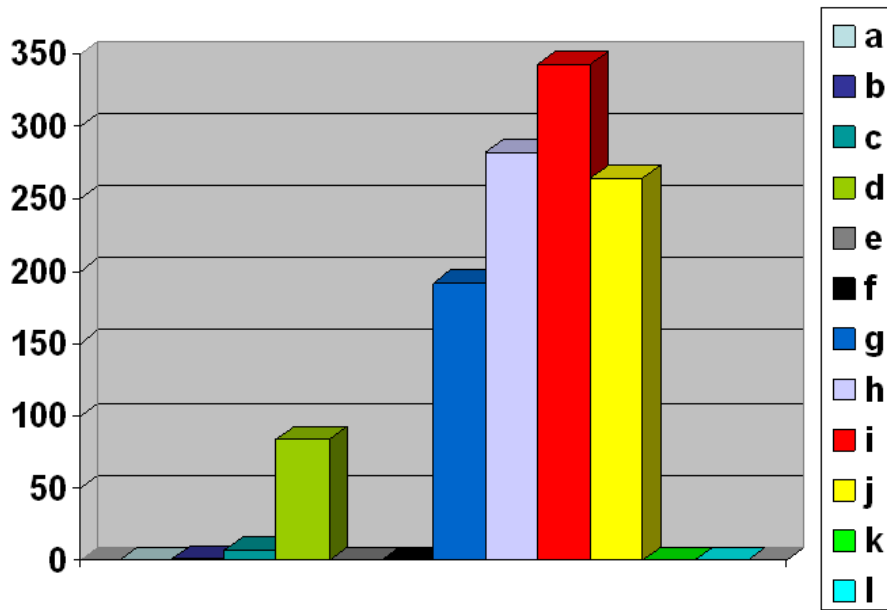
توزیع برجسب‌ها برای چهار تصویر انتخابی در شکل‌های ۳-۱۵ تا ۳-۱۸ نشان داده شده‌است.



(ب)

(الف)

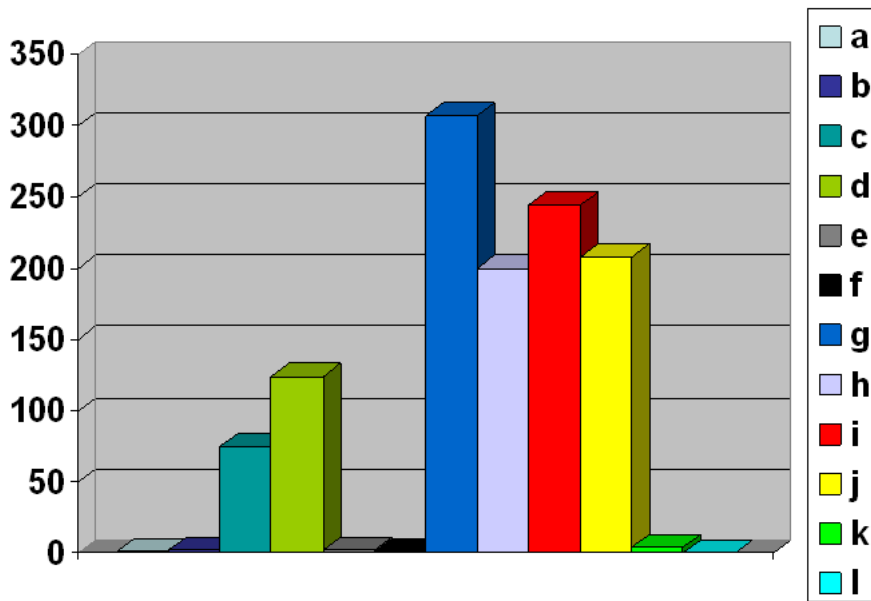
شکل ۳-۱۵: الف) تصویر چشم راست ب) فراوانی برجسب‌ها در تصویر برجسب خورده توپوگرافیکی



(ب)

(الف)

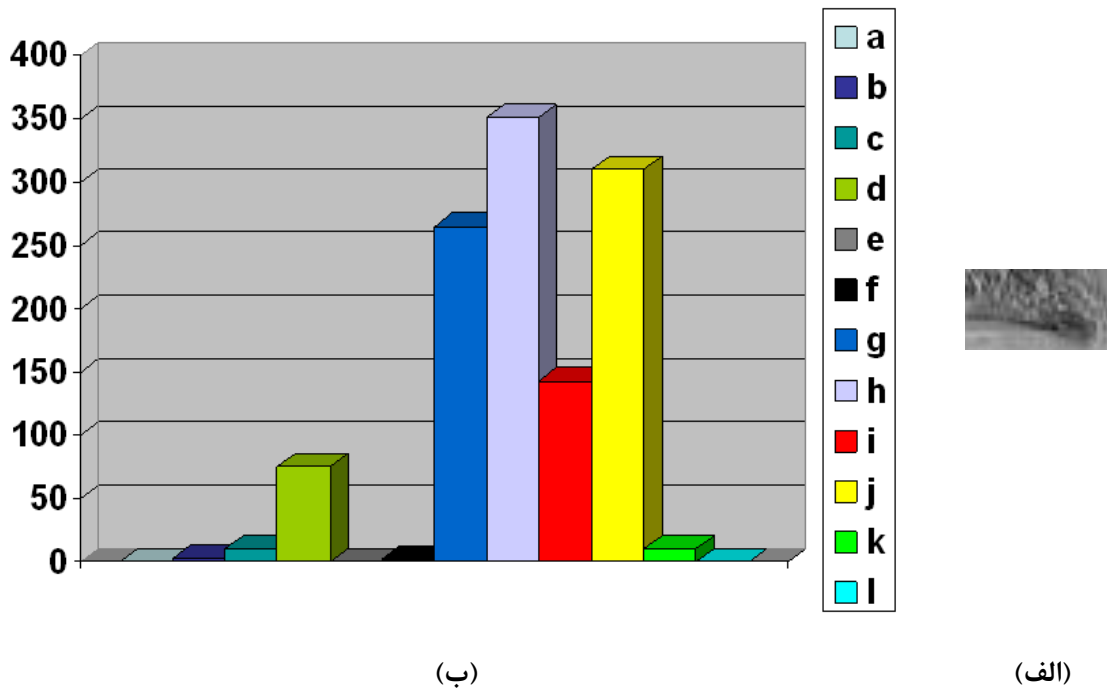
شکل ۳-۱۶: الف) تصویر چشم چپ ب) فراوانی برجسبها در تصویر برجسب خورده توپوگرافیکی



(ب)

(الف)

شکل ۳-۱۷: الف) تصویر غیر چشم ب) فراوانی برجسبها در تصویر برجسب خورده توپوگرافیکی



شکل ۳-۱۸: الف) تصویر غیرچشم ب) فراوانی برجسبها در تصویر برجسب خورده توپوگرافیکی

مقایسه شکل ۳-۱۵ (ب) و شکل ۳-۱۶ (ب) نشان می‌دهد که توزیع برجسب‌های توپوگرافیک در تصاویر چشم مشابه می‌باشد، این حاکی از منحصر بفرد بودن الگوی بدست آمده برای هر کدام از چشم‌ها و متمایز بودن آن از سایر اجزاء چهره می‌باشد.

با تعیین الگوی توپوگرافیک هر ناحیه از تصویر، توزیع برجسبها در آن مشخص می‌گردد. مقایسه توزیع برجسبها در ناحیه آزمون و توزیع قابل انتظار در ناحیه چشم معیار مناسبی برای تشخیص چشم در تصویر چهره می‌باشد [۴۷]. محققین در مرجع [۴۷] به کمک توزیع آماری برجسبها، روشی جهت تشخیص چشم در تصویر چهره ارائه نمودند. در ویژگی‌های بکار رفته، شکل هندسی و پراکندگی پیکسل‌ها در هر ناحیه برجسب خورده نادیده گرفته می‌شود. این ویژگی‌ها فقط فراوانی پیکسل‌ها در هر ناحیه می‌باشند.

در این تحقیق سعی شده است تا با بهره‌گیری از اطلاعات مربوط به پراکندگی پیکسل‌ها در ناحیه مربوط به هر برجسب قدم مؤثری در جهت غنی کردن بردار ویژگی‌ها و افزایش کارایی سیستم تشخیص چشم برداشته شود. تحقیقات ما در این پایان‌نامه نشان می‌دهد که ممان‌های تغییر ناپذیر نسبت به تبدیل شباهتی^۱ کمک شایانی در مقاوم سازی تشخیص چشم دارد.

¹ Similarity Transformation

۳-۲- استخراج ویژگی‌ها با استفاده از ممان‌های ثابت

در بسیاری از کاربردهای شناسایی الگو^۱، تشخیص و بازشناسایی اشیاء و ی‌ا به طور کلی الگوها مقاوم بودن الگوریتم نسبت به جابجائی مکان الگو در تصاویر و چرخش آن از اهمیت خاصی برخوردار بوده است. می‌توان روش‌هایی که در جهت تغییر ناپذیر کردن ویژگی‌ها به تبدیل‌های هندسی پیشنهاد می‌شوند را به دو دسته اساسی تقسیم کرد:

۱. الگوریتم‌های مبتنی بر مجموعه‌ای از ویژگی‌ها^۲

۲. روش‌های آماری که خود شامل تئوری تصمیم و روش شبکه‌های تصادفی می‌شوند.

روش‌های مبتنی بر مجموعه ویژگی، در صورتی نتایج مطلوبی ارائه می‌دهند که برای الگوهای خاصی طراحی و تدوین شده باشند. شرط مذکور مؤید این مطلب است که این روش، عمومی نیست. البته این روش از دیدگاه تئوری، روشی مستقل از تغییرات مکان، ابعاد و چرخش می‌باشد. اما بزرگترین محدودیت آن در این است که فقط برای کاربرد خاص و نمونه‌های ورودی خاص قابل اعمال است و در صورتی که نمونه‌های ورودی تغییر کند، الگوریتم دچار خطا می‌گردد. از سوی دیگر روش آماری این محدودیت را نداشته و قادر است بر روی مجموعه نمونه‌های جدید ورودی نیز اعمال گردد. اما متأسفانه قدرت این روش نسبت به تغییرات مکان، ابعاد و چرخش به خوبی روش قبل نیست.

Hu [۴۸] روشی بر اساس ممان‌های آماری تغییر ناپذیر^۳ و کاربرد آنها در پردازش تصویر و بازشناسایی الگو ارائه داد. نتایج حاصل حاکی از آن بود که الگوریتم کاملاً نسبت به جابجائی مکان تصاویر و چرخش مقاوم بوده و از طرف دیگر قابلیت یادگیری بر اساس مجموعه الگوهای آموزشی را نیز دارد. علاوه بر زمینه‌های مذکور مفهوم ممان به طور وسیعی در مکانیک کلاسیک و تئوری آماری، مورد استفاده قرار می‌گیرد. از جمله مفاهیمی که در رابطه با ممان‌ها مطرح می‌گردد، هنجارسازی اندازه^۴ و محورهای اصلی^۵ می‌باشند. در حوزه بازشناسایی الگو، هنجارسازی اندازه و مرکز جرم به عنوان پیش‌پردازش در نظر گرفته شده‌اند [۴۹][۵۰]. در مرجع [۵۱] هنجارسازی چرخش نیز مورد بررسی قرار گرفته است. Hu [۴۸] نیز روشی مستقل از چرخش در استفاده از ممان‌های آماری مطلق و ی‌ا عمود برهم ارائه می‌دهد. در بخش بعد به تعریف مفاهیم اشاره کرده و سپس توصیف آماری نواحی به کمک ممان‌های تغییر ناپذیر را توضیح خواهیم داد.

¹ Pattern Recognition

² Property list

³ Moment invariant

⁴ Size normalization

⁵ Principal axes

۳-۲-۱- ممان‌ها و ثابت‌های جبری^۱

در روش Hu [۴۸]، ممانهای تابع توزیع چگالی $\rho(x, y)$ از مرتبه $(p+q)$ با انتگرال ری‌مان^۲ بصورت زیر محاسبه می‌گردند:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q \rho(x, y) dx dy, \quad p, q = 0, 1, 2, \dots \quad ۲۶-۳$$

تابع $\rho(x, y)$ به صورت تابعی تکه‌ای پیوسته فرض شده که فقط در ناحیه محدودی در صفحه xy غیر صفر است. در این صورت ممان تمام مرتبه‌ها بر اساس تعریف بالا وجود دارد و تئوری یکتایی زیر نیز برقرار است:

هر رشته دوتایی از ممانها $\{m_{pq}\}$ به طوری که از روی $\rho(x, y)$ تعریف می‌گردد؛ عکس این قضیه نیز برقرار است، یعنی $\rho(x, y)$ به تنهایی از روی $\{m_{pq}\}$ قابل استخراج است.

۳-۲-۲- تابع مشخصه و تابع مولد ممان^۳

تابع مشخصه^۴ و تابع مولد ممان^۵ برای توزیع $\rho(x, y)$ به ترتیب به صورت زیر تعریف می‌گردند:

$$\phi(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(iux + ivy) \rho(x, y) dx dy, \quad ۲۷-۳$$

$$M(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(ux + vy) \rho(x, y) dx dy. \quad ۲۸-۳$$

در توابع فوق u و v حقیقی می‌باشند. اگر ممانهای همه مراتب موجود باشند دو تابع فوق با استفاده از سری‌های توانی برحسب ممان‌های $\rho(x, y)$ به صورت زیر بدست می‌آیند [۴۸]:

$$\phi(u, v) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} m_{pq} \frac{(iu)^p}{p!} \frac{(iv)^q}{q!}, \quad ۲۹-۳$$

$$M(u, v) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} m_{pq} \frac{u^p}{p!} \frac{v^q}{q!}. \quad ۳۰-۳$$

این دو تابع به طور گسترده در تئوری آمار استفاده می‌شوند. اگر تابع مشخصه $\phi(u, v)$ که اساساً تبدیلی فوری^۵ تابع $\rho(x, y)$ است معلوم باشد، $\rho(x, y)$ با توجه به معکوس تبدیلی فوری^۵ بصورت زیر بدست می‌آید:

¹ Algebraic Invariants

² Riemann

³ Characteristic Function and Moment Generating Function

⁴ Characteristic function

⁵ Moment generating function

$$\rho(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(-iux - ivy) \phi(x, y) dudv. \quad ۳۱-۳$$

۳-۲-۳ ممان مرکزی^۱

ممان مرکزی مرتبه^۲ p و q تابع $\rho(x, y)$ که با μ_{pq} نمایش داده می شود بصورت زیر تعریف می گردد:

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q \rho(x, y) d(x - \bar{x}) d(y - \bar{y}), \quad ۳۲-۳$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

انتقال محورهای مختصات به صورت زیر ممان های مرکزی را تغییر نمی دهد [۴۸].

$$\begin{aligned} x' &= x + \alpha, \\ y' &= y + \beta, \quad \alpha, \beta \rightarrow const. \end{aligned}$$

با توجه به رابطه^{۳-۳۲} به راحتی می توان همه ممان های یک تابع را برحسب ممان مرکزی آن بی ان نمود. برای چهار ممان اول تابع $\rho(x, y)$ داریم:

$$\begin{aligned} \mu_{00} &= m_{00} = \mu \\ \mu_{10} &= \mu_{01} = 0 \\ \mu_{20} &= m_{20} - \mu \bar{x}^2 \\ \mu_{11} &= m_{11} - \mu \bar{x} \bar{y} \\ \mu_{02} &= m_{02} - \mu \bar{y}^2 \\ \mu_{30} &= m_{30} - 3m_{20} \bar{x} + 2\mu \bar{x}^3 \\ \mu_{21} &= m_{21} - m_{20} \bar{y} - 2m_{11} \bar{x} + 2\mu \bar{x}^2 \bar{y} \\ \mu_{12} &= m_{12} - m_{02} \bar{x} - 2m_{11} \bar{y} + 2\mu \bar{x} \bar{y}^2 \\ \mu_{03} &= m_{03} - 3m_{02} \bar{y} + 2\mu \bar{y}^3 \end{aligned} \quad ۳۳-۳$$

توابع ممان تغییر ناپذیر متعامد (ضمیمه الف) به آسانی برای تشخیص ص الكوهای مستقل از چرخش به کار گرفته می شوند. اگر این توابع با توابع تشابه ممان مرکزی ترکیب گردند، فرایند تشخیص ص الكو نسبت به ابعاد و مکان مستقل خواهد شد.

بر این اساس [۴۸] Hu مجموعه ای از هفت تابع که ترکیبی از ممان ها می باشند و توانایی توصیف و نمایش هر شیء دلخواه را دارند ارائه نمود:

^۱ Central Moments

$$\begin{aligned} \phi_1 &= \mu_{20} + \mu_{02} \\ \phi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\ \phi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\ \phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 & ۳۴-۳ \\ \phi_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\ &+ (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\ \phi_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \end{aligned}$$

و همچنین یک ثابت متعامد مایل:

$$\begin{aligned} \phi_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] & ۳۵-۳ \\ &- (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \end{aligned}$$

ثابت اخیری برای تشخیص تصاویر آینه ای مناسب می باشد.

این روش علاوه بر قابلیت اعمال روی اهداف مستقل از چرخش، ابعاد و مکان برای موارد مستقل از نگاشت موازی^۱ هم کاربرد دارد.

ممانهای عنوان شده در بالا از روی تصویر باینری تقطیع شده یا تصویر مربوط به مرزهای شیء مورد نظر محاسبه می گردند. Jiang و همکارش [۵۲] نشان دادند که هر دوی این روشها نتایج یکسانی ارائه می دهند در حالی که Dudani و همکارانش [۵۳] بیان می دارند که محاسبه از روی تصویر باینری نسبت به نویز مقاومتر است.

۳-۲-۴ - محاسبه بردار ویژگیها

همانطور که بخش ۳-۱ بیان شد، ساختار توپوگرافیک تصویر چشم از الگوی خاصی پیروی می کند که این الگو در تمام تصاویر برچسب خورده توپوگرافیکی چشم ملاحظه می شود. در روشی که توسط Wang و همکارش [۴۷] ارائه شد فراوانی هر یک از برچسبهای توپوگرافیک در ناحیه چشم به عنوان ویژگی برای تشخیص این عضو استفاده شد. همانطور که اشاره شد جهت غنی سازی ویژگیهای توپوگرافیکی از توابع ممان تغییر ناپذیر برای توصیف نواحی بدست آمده برای برچسبهای توپوگرافیکی در ناحیه چشم استفاده می کنیم.

در صورتیکه تصویر برچسب خورده توپوگرافیکی به زیر تصویرها تجزیه شود بطوریکه در هر زیر تصویر تنها ناحیه مربوط به یک برچسب از برچسبهای دوازده گانه توپوگرافیکی ظاهر گردد، بخشهای مختلف چشم از یکدیگر جدا می شوند. بدیهی است این زیر تصویرها در اثر انتقال، چرخش یا تغییر ابعاد، شکل خود را حفظ می کنند. همانطوریکه در شکل های ۳-۱۱ تا ۳-۱۴ ملاحظه می گردد، ارزش پیکسلها در هر زیر تصویر ۰ یا ۱ می باشد.

¹ Parallel projection

به منظور استخراج ویژگی‌های تغییر ناپذیر در اثر تبدیل هندسی تشابهی از هفت تابع زیر که در روش Hu [۴۸] تعریف شده‌است، استفاده می‌کنیم:

$$\begin{aligned}
 \phi_1 &= \mu_{20} + \mu_{02} \\
 \phi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\
 \phi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\
 \phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \\
 \phi_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
 &\quad + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\
 \phi_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \\
 \phi_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
 &\quad - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]
 \end{aligned}
 \tag{۳۶-۳}$$

با توجه به نتایج مرحله قبل که حاکی از وجود نواحی مهمی از برجسب‌های توپوگرافیک (ک، ی، ط، ح، د) در تصاویر کلاس چشم بود (شکل ۳-۶)، فقط ویژگی‌های آماری مربوط به نواحی هر کدام از این پنج برجسب جهت توصیف تصویر چشمها بکار گرفته می‌شود. همانطوری که قبلاً اشاره شد از ناحیه مربوط به هر برجسب، هفت ویژگی آماری تغییری ناپذیری استخراج می‌شود و در مجموع به کمک ۳۵ ویژگی آماری توصیف مناسبی برای تصویر چشم ایجاد می‌کنیم. بردار ویژگی توصیفگر هر تصویر، یک بردار ستونی X (۳۷-۳) با $7d$ مؤلفه است که d تعداد برجسب‌های توپوگرافیکی بکار رفته در توصیف تصویر می‌باشد ($d=5$).

$$X = \begin{bmatrix} \phi_1^1 \\ \phi_2^1 \\ \vdots \\ \phi_1^d \\ \vdots \\ \phi_7^d \end{bmatrix}
 \tag{۳۷-۳}$$

ϕ_1^j ام ویژگی مربوط به تابع ممان تغییر ناپذیر ϕ_1 (رابطه ۳-۳۶) مربوط به ناحیه برجسب ϕ_1 و ϕ_2^j ام ویژگی مربوط به تابع ممان ϕ_2 از همان ناحیه برجسب خورده می‌باشد.

۳-۲-۵- تخمین پارامترهای مدل آماری برای کلاس‌های چشم راست، چشم چپ و نواحی غیر چشم

پس از محاسبه بردار ویژگی‌ها برای نمونه‌های آموزشی در هر کلاس، هر نمونه نقطه‌ای در فضای n ($n=7d$) بعدی می‌باشد. با در نظر گرفتن توزیع نرمال برای ویژگی‌های استخراج شده از هر کلاس بردار میانگین و ماتریس کواریانس متناظر با کلاس به کمک داده‌های آموزشی تخمین زده می‌شوند.

توزیع نرمال چند متغیره n بعدی توسط رابطه زیر بیان می‌شود:

$$P(\vec{X}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\vec{X} - \vec{\mu})^T \Sigma^{-1} (\vec{X} - \vec{\mu})\right) \quad 38-3$$

در این رابطه $\vec{X} = (x_1, x_2, \dots, x_n)^T$ بردار ویژگی، $\vec{\mu} = (\mu_1, \mu_2, \dots, \mu_d)^T$ بردار میانگین و Σ ماتریس کواریانس هستند. ابعاد بردار ویژگی و بردار میانگین برابر با $(n \times 1)$ و ابعاد ماتریس کواریانس $(n \times n)$ می‌باشد.

بردار میانگین و ماتریس کواریانس در روش تخمین ML (Maximum Likelihood) از روی m نمونه آموزشی بصورت زیر بدست می‌آید:

$$\vec{\mu} = \frac{1}{n} \sum_{k=1}^m \vec{X}_k \quad 39-3$$

$$\Sigma = \frac{1}{m} \sum_{k=1}^m (\vec{X}_k - \vec{\mu})(\vec{X}_k - \vec{\mu})^T \quad 40-3$$

در این تحقیق نمونه‌های آموزشی در سه گروه چشم راست، چشم چپ و غیر چشم در نظر گرفته می‌شوند. نکته قابل ذکر در رابطه با گروه غیر چشم این است که تصاویر این گروه از مناطق اطراف مینیمم‌های محلی (غیر از مردمک) در تصاویر چهره انتخاب می‌شوند.

3-3- تعیین مکان چشم‌ها در تصویر آزمون

در این تحقیق الگوریتم مربوط به تعیین مکان چشم‌ها در تصویر آزمون به مراحل زیر تقسیم می‌شود:

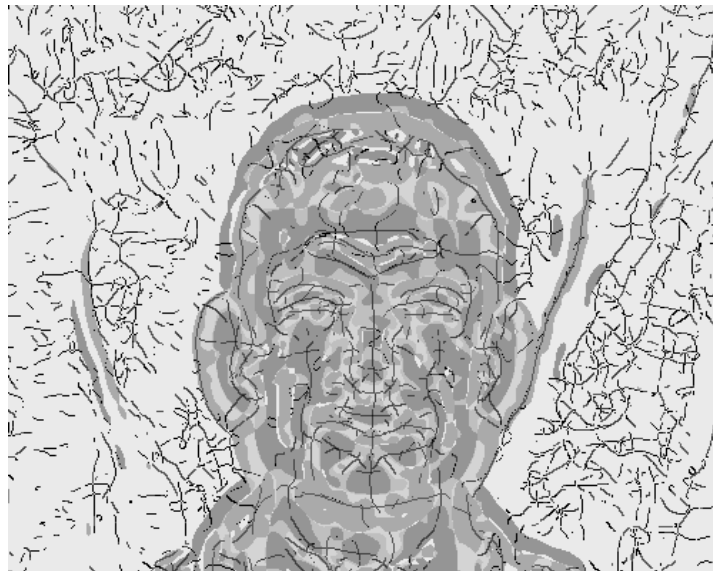
۱. بدست آوردن تصویر برچسب خورده توپوگرافیکی از روی تصویر ورودی
۲. استخراج نقاط کاندیدای حضور مردمک (برچسب‌های چاله)
۳. گروه بندی کاندیدها و حذف کاندیدهای غیر ممکن برای چشم
۴. انتخاب محل اصلی مردمک و جداسازی ناحیه چشم از تصویر ورودی

3-3-1- تعیین ساختار توپوگرافیک تصویر ورودی

با توجه به اینکه ویژگی‌های ظاهر شده در حوزه توپوگرافیک اساس تعیین مدل چشم در چهره بشمار می‌روند، اولین قدم در الگوریتم تشخیص چشم نیز تعیین برچسب‌های توپوگرافیکی تصویر آزمون می‌باشد. شکل 3-19 تصویری از چهره را به همراه تصویر برچسب خورده توپوگرافیکی آن نمایش می‌دهد.



(الف) تصویر ورودی

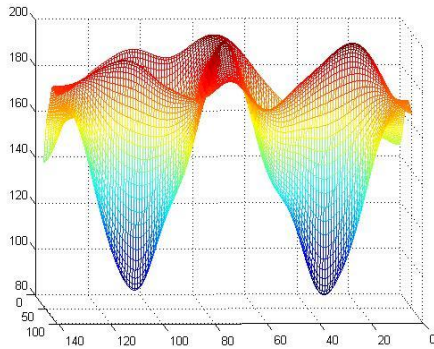


(ب) تصویر برچسب خورده توپوگرافیکی

شکل ۳-۱۹: تعیین ساختار توپوگرافیک تصویر آزمون

۳-۳-۲- تعیین نقاط کاندیدای حضور مردمک

برای تعیین نقاط کاندیدای مردمک از وجود برچسب چاله (pit)، در تصویر توپوگرافیکی چهره استفاده می‌کنیم. یکی از محل‌های حضور این برچسب مردمک می‌باشد. شکل ۳-۲۰، تصویر قسمتی از چهره شامل چشم‌ها و ابروها را به همراه رویه متناظر با آن نشان می‌دهد. مینیمم‌های محلی که در مرکز مردمک‌ها اتفاق می‌افتد در شکل ۳-۲۰ (ب) کاملاً مشخص است.

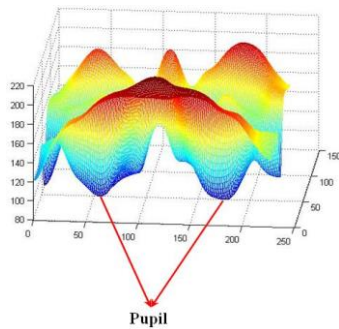


(ب): رویه تصویر الف

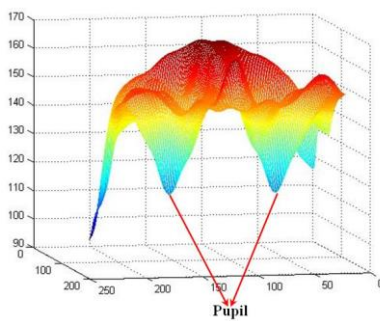
(الف): تصویر چهره در ناحیه چشمها و ابروها

شکل ۳-۲۰: ویژگی توپوگرافیکی مردمک

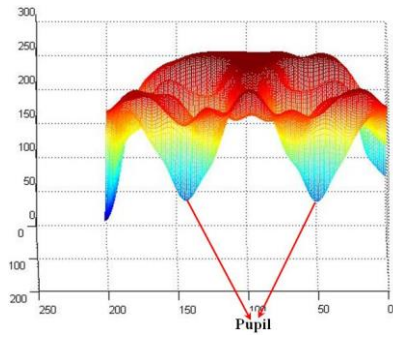
با توجه به اینکه ویژگی توپوگرافیکی اشاره شده برای مردمک به حضور عوامل خارجی مثل عینک بستگی ندارد و در شرایط مختلف تصویر برداری باقی می ماند، این ویژگی معیار مناسبی برای انتخاب نقاط کاندیدای مردمک می باشد. شکل ۳-۲۱ رویه مربوط به چهره را در ناحیه چشم برای افراد مختلف و در شرایط با عینک و بدون عینک و نورپردازی های مختلف نشان می دهد. در تمام این حالات وجود مینیمم های محلی در مرکز مردمک ملاحظه می گردد.



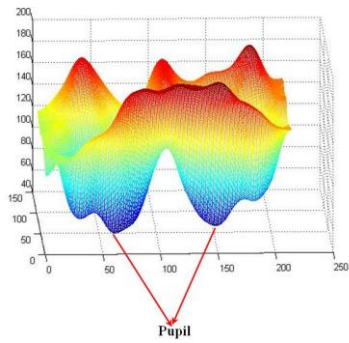
شکل ۳-۲۱(الف): قرار گرفتن عینک روی چشمها



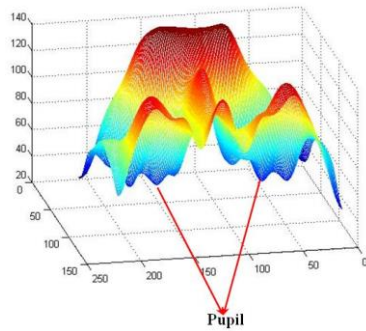
شکل ۳-۲۱(ب): کنتراست پایین



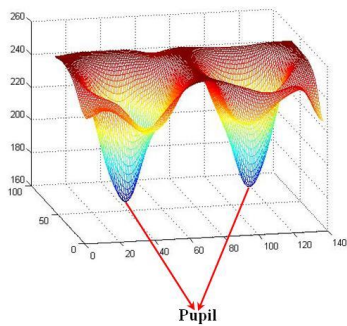
شکل ۳-۲۱(ج) : کنتراست بالا



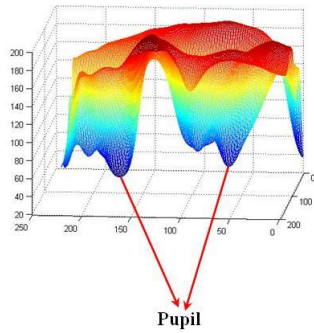
شکل ۳-۲۱(د): انقباض اطراف چشمها



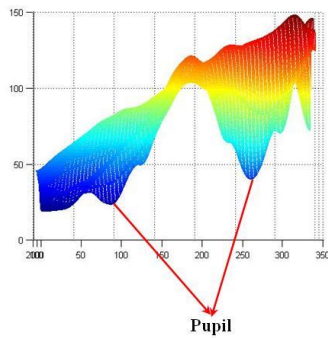
شکل ۳-۲۱(ه) : روشنایی پایین



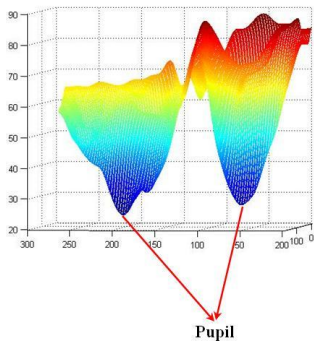
شکل ۳-۲۱(و) : روشنایی بالا



شکل ۳-۲۱(ی): چرخش سر



شکل ۳-۲۱(ن): تابش نور از یک جهت



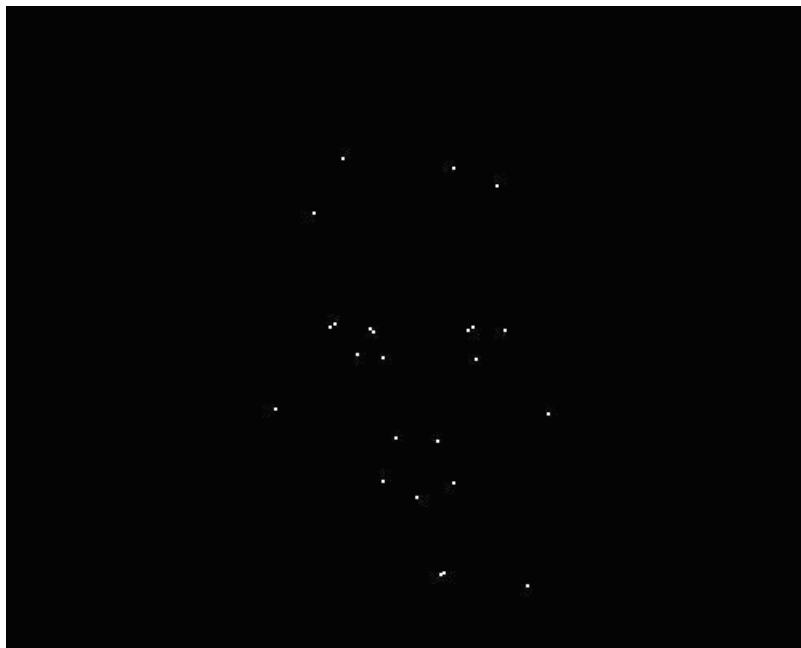
شکل ۳-۲۱(م): تنگ شدن چشم‌ها

شکل ۳-۲۱: بررسی نقطه کاندید مردمک در شرایط مختلف

در این مرحله پس از تعیین ساختار توپوگرافیک تصویر، کلیه نقاطی که برچسب چاله (مینیمم محلی) به آنها اختصاص داده شده است در لیست نقاط کاندید مردمک قرار می‌گیرند. شکل ۳-۲۲ تصویری از نقاط کاندید را در تصویر یک چهره نشان می‌دهد.



(الف): تصویر چهره



(ب): نقاط کاندید مردمک در تصویر

شکل ۳-۲۲: تعیین نقاط کاندیدای مردمک

۳-۳-۳ گروه بندی کاندیدها و حذف تعدادی از آنها

گروه بندی و حذف کاندیدها در ۳ مرحله انجام می شود. این مراحل به ترتیب ذیل می باشد:
۱. حذف چاله هایی که در همسایگی آنها برجسب سرایشی بدون انحناء^۱ وجود دارد.

^۱ Slop hill

۲. حذف کاندیدهایی که در معیار هندسی حاکم بر چشم‌ها صدق نمی‌کنند.
۳. حذف کاندیدهای چشم راست و چپ که در ناحیه بین آنها برچسب ناودانی بیرون^۱ وجود ندارد.
- همانطور که در بخش ۳-۱-۳ گفته شد، برچسب سرایشی بدون انحناء در ناحیه چشم‌ها وجود ندارد. شکل ۳-۲۳ یک تصویر آزمون را به همراه توزیع برچسب سرایشی بدون انحناء را در تصویر توپوگرافیکی آن نشان می‌دهد.



شکل ۳-۲۳ (الف): تصویر آزمون



(ب) : توزیع برچسب سرایشی بدون انحنا در تصویر توپوگرافیکی

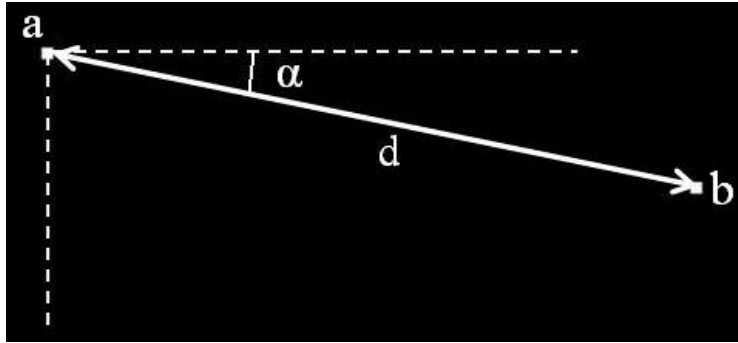
شکل ۳-۲۳: استفاده از برچسب سرایشی بدون انحنا در دسته‌بندی و حذف کاندیدهای نادرست

بررسی توزیع برچسب سرایشی بدون انحنا در تصاویر غیر چشم نشان داد که فراوانی این برچسب در موها، ریش، پوست و پس زمینه بسیار زیاد است. در صورتیکه این برچسب در ناحیه چشم‌ها ظاهر نمی‌شود. با توجه به اینکه چاله‌های زیادی در موها و ریش افراد وجود دارند، می‌توان با حذف چاله‌هایی که در همسایگی آنها برچسب سرایشی بدون انحنا وجود دارد، تعداد زیادی از کاندیدهای نادرست را حذف کرد. در الگوریتم پیشنهادی چنانچه در همسایگی چاله برچسب سرایشی بدون انحنا وجود داشته باشد، چاله مذکور از لیست کاندیدهای مردمک حذف می‌شود. دلیل اینکه مابقی برچسب‌هایی که در ناحیه چشم حضور ندارند در این طبقه‌بندی مورد استفاده قرار نمی‌گیرند، این است که توزیع این برچسب‌ها در نواحی دیگر چهره نیز ضعیف می‌باشد (شکل‌های ۳-۱۵ تا ۳-۱۸).

کاملاً واضح است که تعداد زیادی از نقاط کاندید را می‌توان با استفاده از معیارهای هندسی در مراحل ابتدایی حذف نمود. با توجه به اینکه این نقاط بصورت دوه‌دو بعنوان مردمک چشم راست و مردمک چشم چپ مورد آزمایش قرار می‌گیرند، معیارهایی که برای رد کاندیدهای نادرست مورد استفاده قرار می‌دهیم فاصله بین زوج نقطه کاندید و زاویه خط واصل دو نقطه با محور افقی است.

در شکل ۳-۲۴، فاصله بین دو نقطه a و b که به ترتیب کاندیدهای مردمک چشم‌های راست و چپ بشمار می‌آیند با d نشان داده شده‌است. زاویه بین خط واصل دو نقطه a و b و محور افقی تصویر نیز با α مشخص شده‌است.

اگر مختصات نقاط a و b به ترتیب با (x_1, y_1) و (x_2, y_2) بیان شود، d و α با روابط ۴۱-۳ و ۴۲-۳ تعیین می گردند.



شکل ۳-۲۴: خصوصیات هندسی مربوط به نقاط کاندید

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad ۴۱-۳$$

$$\alpha = \tan^{-1} \left(\frac{x_2 - x_1}{y_2 - y_1} \right) \quad ۴۲-۳$$

در این مرحله ابتدا کلیه کاندیدها شماره گذاری می شوند، شماره گذاری به ترتیب از گوشه بالای سمت چپ تصویر تا پایین در جهت افزایش سطرها و ستونها ادامه پیدا می کند. مجموعه C از نقاط شماره گذاری شده بصورت $C = \{c_1, \dots, c_L, \dots, c_K\}$ تشکیل می شود که در این مجموعه، c_L ، L امین کاندید می باشد، سپس برای هر کاندید، دو فهرست از فاصله و زاویه بین خط واصل آن تا دیگر کاندیدها و محور افقی تهیه می شود. برای جلوگیری از تکرار در انتخاب زوج کاندیدها، نقاطی در تهیه این فهرست دخالت دارند که اندیس اختصاص داده شده به آنها بزرگتر از اندیس کاندید مورد نظر باشد. دو مجموعه مربوط به فواصل و زاویه های بدست آمده برای کاندید c_L را بصورت زیر داریم:

$$D_L = \{d_{L,L+1}, d_{L,L+2}, \dots, d_{L,L+K}\} \quad ۴۳-۳$$

$$\theta_L = \{\alpha_{L,L+1}, \alpha_{L,L+2}, \dots, \alpha_{L,L+k}\}$$

با تشکیل دو مجموعه D_L و θ_L برای کلیه کاندیدها، فاصله و زاویه بین تمام زوج هایی که کاندید مردمک چشم راست و چپ هستند مشخص می گردد.

با توجه به رنج تغییرات فاصله دوربین تا چهره، نسبت ابعاد چهره به کل تصویر در محدوده مشخصی خواهد بود. لذا فاصله مردمک چشمها از یکدیگر می بایست در رنج مشخصی نسبت به ابعاد تصویر باشد. این یکی از معیارهای هندسی است که برای فیلتر کردن کاندیدهای نادرست بکار می بریم. از طرف دیگر چرخش سر در صفحه عمود بر محور اپتیکی دوربین در شرایط عادی محدود است، این چرخش را می توان به کمک زاویه ای که خط واصل مردمکها با محور افقی تولید می کند، بدست آورد. معیار دوم جهت حذف کاندیدهای زوج مردمک لحاظ نمودن این محدودیت زاویه می باشد.

در این تحقیق آستانه مربوط به فاصله و زاویه بین نقاط با توجه به ابعاد تصویر ورودی، آگاهی قبلی در رابطه با فاصله شخص و دوربین و ماکزیمم چرخش سطحی و عمقی سر تعیین شد.

در سومین مرحله از حذف کاندیدهای نادرست مجدداً اطلاعات مربوط به برچسب‌های توپوگرافیکی مورد استفاده قرار می‌گیرند. بررسی تصویر برچسب خورده توپوگرافیکی نشان می‌دهد که یکی از برچسب‌های اختصاص داده شده به بینی، برچسب ناودانی بیرون می‌باشد. حضور این برچسب در محل بینی ویژگی بسیار قدرتمندی در تصویر توپوگرافیکی محسوب می‌گردد. شکل ۳-۲۵ توزیع این برچسب را برای تصویر آزمون نشان داده شده در شکل ۳-۲۳(الف) نشان می‌دهد. در چهره، بینی در ناحیه بین دو چشم قرار دارد، بنابراین وجود برچسب ناودانی بیرون بین چاله‌های مربوط به مردمک‌ها در تصویر برچسب خورده توپوگرافیکی الزامی می‌باشد. در الگوریتم پیشنهادی از این معیار برای حذف کاندیدهای نادرست استفاده شد. به این منظور برای کلیه زوج کاندیدهایی که به این مرحله وارد شدند یک ناحیه جستجو تعریف کردیم. این ناحیه مستطیل شکل دارای ابعاد $0.5d \times 0.25d$ (d فاصله بین دو مردمک می‌باشد) می‌باشد و مرکز آن روی نقطه میانی خط واصل دو چاله قرار گرفته است. در صورتیکه در این ناحیه برچسب ناودانی بیرون وجود نداشته باشد، چاله‌های مذکور از لیست کاندیدهای مردمک حذف می‌شوند.



شکل ۳-۲۵: توزیع برچسب ناودانی بیرون در تصویر برچسب خورده توپوگرافیکی شکل ۳-۲۳(الف)

با حذف کاندیدهای ناممکن به شیوه گفته شده، تعداد زوج کاندیدهای مردمک که به مرحله نهایی وارد می‌شوند به میزان قابل توجهی کاهش می‌یابد. ورودی مرحله بعد الگوریتم زوج‌های مرتب متشکل از کاندیدهای مراکز مردمک می‌باشند.

۳-۳-۴ - تعیین محل مردمک و جداسازی ناحیه چشم از تصویر ورودی

همانطور که قبلاً گفته شد هر نمونه آموزشی بصورت یک نقطه در فضای n بعدی (n تعداد ویژگی‌های استخراج شده از تصویر می‌باشد) مشخص می‌شود. نمونه‌های مربوط به هر یک از سه کلاس چشم راست، چشم چپ و غیر چشم تشکیل توده‌هایی با مراکز تجمع جدا از هم بنام کلاستر^۱ می‌دهند. با توجه به شباهت زیادی که بین دو کلاس چشم راست و چشم چپ وجود دارد مراکز کلاسترهای مربوط به این دو گروه نسبت به کلاستر گروه سوم نزدیک می‌باشند. در پروسه کلاسه‌بندی یک نمونه آزمایشی وارد فضای n بعدی شده و با تعیین فاصله این نمونه با مراکز سه کلاستر می‌توان تعلق آن را به یکی از سه کلاس چشم راست، چشم چپ و غیر چشم تعیین نمود.

کلاسه‌بندی مطابق با مراحل نشان داده شده در بلوک دیاگرام شکل ۳-۲۶ انجام می‌شود. ورودی‌های این مرحله جفت نقاطی از کاندیدهای مردمک هستند که از مرحله قبل بدست آمده‌اند. فاصله بین زوج کاندید (d) و زاویه خط واصل دو نقطه و محور افقی (α) نیز از مرحله قبل مشخص می‌باشد.

همانطور که در بلوک دیاگرام شکل ۳-۲۶ دیده می‌شود با ورود زوج نقطه مورد آزمایش که a و b نامیده شده‌اند، مراحل زیر به ترتیب انجام می‌شوند:

۱. در اولین مرحله یک ماسک شامل دو ناحیه مستطیل شکل به مرکز a و b و در راستای خط واصل دو نقطه کاندید ایجاد می‌شود. ابعاد این مستطیل‌ها $0.6d \times 0.3d$ می‌باشد.

۲. ماسک تهیه شده بر روی تصویر مربوط به ساختار توپوگرافیک چهره قرار گرفته و ناحیه اطراف نقاط کاندید از تصویر برچسب خورده جدا می‌شوند. دو ناحیه جدا شده در این مرحله کاندیدهای چشم راست و چشم چپ می‌باشند.

۳. بردار ویژگی برای هر یک از دو ناحیه استخراج شده، تشکیل می‌گردد.

۴. در این مرحله از یک کلاسه‌بند بیزین^۲ برای کلاسه‌بندی نواحی استخراج شده استفاده می‌شود. فاصله ماهالانوبیس^۳ بین بردار ویژگی هر یک از دو ناحیه با مراکز کلاسترهای هر یک از سه گروه محاسبه می‌شود. این فاصله توسط رابطه زیر محاسبه می‌گردد.

$$D^2 = (\bar{x} - \bar{\mu})^T \Sigma^{-1} (\bar{x} - \bar{\mu}) \quad ۴۴-۳$$

در رابطه بالا $\bar{\mu}$ و Σ به ترتیب بردار میانگین و ماتریس کواریانس مربوط به کلاستری هستند که فاصله بردار ویژگی آزمون \bar{x} از آن محاسبه می‌شود.

¹ Cluster

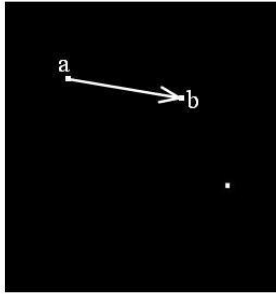
² Bayesian

³ Mahalanobis

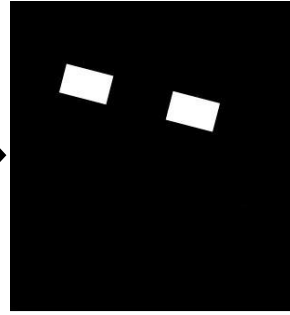
- برای تعیین تعلق هر یک از نواحی به مدل مربوطه‌شان، فاصله ماہالانوبیس بین بردار ویژگی ناحیہ کاندید چشم راست را با مرکز کلاستر مربوط به مدل چشم راست و مرکز کلاستر مربوط به مدل غیر چشم محاسبه می‌کنیم. ناحیہ مورد آزمون متعلق به کلاستری است که با آن فاصله کمتری دارد. عمل کلاسه‌بندی برای ناحیہ چشم چپ نیز انجام می‌شود.

با تکرار مراحل ۱ تا ۴ از رویہ فوق برای هر یک از زوج کاندیدهای وارد شده به این پروسه میزان تعلق عناصر زوج به سه گروه چشم راست، چشم چپ یا غیر چشم مشخص می‌شود. در این صورت کاندید به کلاستری متعلق می‌باشد که فاصله ماہالانوبیس کمتری تا مدل کلاستر مربوطه داشته باشد. با خاتمہ زوج کاندیدهای مورد آزمایش، از بین زوج‌هایی که به گروه چشم راست و چپ نسبت داده می‌شوند، جفتی که حاصلجمع دو فاصله ماہالانوبیس محاسبه شده از مراکز کلاسترهای چشم چپ و راست برای آن مینیمم مقدار را در بین کلیہ نمونه‌ها داشته باشد بعنوان چشم راست و چپ در تصویر آزمون انتخاب می‌شوند.

به منظور مشاهده عملکرد الگوریتم تحت شرایط مختلف، الگوریتم پیشنهادی با تصاویر نشان داده شده در شکل ۲۱-۳ مورد آزمایش قرار گرفت. نتایج حاصله در شکل ۲۷-۳ نشان داده شده است. تصویر برچسب خورده توپوگرافیکی، نقاط تعیین شده بعنوان کاندید مردمک، توزیع برچسب ناودانی بیرون و توزیع برچسب سرایشی بدون انحناء نیز در این شکل دیده می‌شوند. در تصویر مربوط به نقاط کاندید محل صحیح مردمک با دایره‌ای در اطراف آن مشخص شده است. مقایسہ تصاویر برچسب خورده توپوگرافیکی برای حالات نشان داده شده حاکی از مقاومت بالای ویژگی‌های توپوگرافیکی در اثر تغییر عوامل مختلف می‌باشد.



جداسازی یک ناحیه مستطیلی
 شکل در اطراف دو کاندید مردمک
 در امتداد خط واصل دو نقطه و
 ایجاد یک ماسک



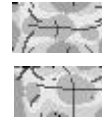
$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$\alpha = \tan^{-1} \left(\frac{x_2 - x_1}{y_2 - y_1} \right)$$

قرار دادن ماسک
 بر روی تصویر
 نگاشت یافته به
 حوزه توپوگرافیک
 و استخراج دو بخش
 کاندید چشم

$$\bar{x}_R = \begin{bmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{bmatrix} \quad \bar{x}_L = \begin{bmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{bmatrix}$$

محاسبه بردار ویژگی برای دو
 منطقه استخراج شده



محاسبه فاصله
 ماهالانوبیس بین X_R با
 کلاس چشم راست و
 غیر چشم و X_L با چشم
 چپ و غیر چشم

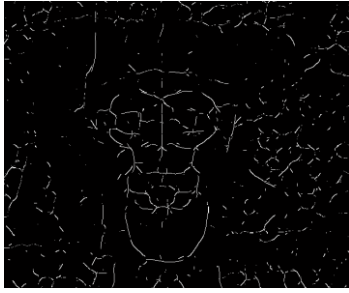
X_R متعلق به کلاس چشم راست یا غیر چشم؟

X_L

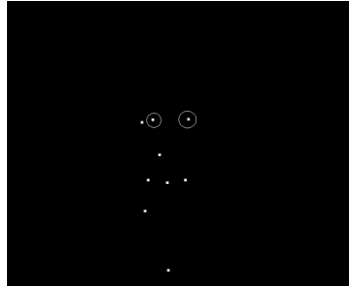
انتخاب زوجی با کمترین فاصله نسبت
 به کلاس چشم راست و چپ و
 بیشترین فاصله با کلاس غیر چشم



شکل ۳-۲۶: مراحل کلاسه‌بندی مناطق کاندید چشم در تصویر



توزیع برجسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برجسب خورده توپوگرافیکی



پاسخ الگوریتم

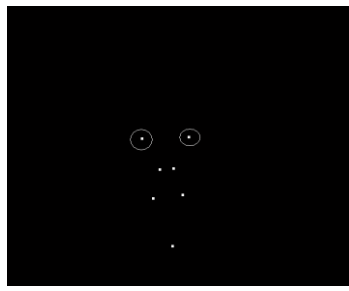


توزیع برجسب سراسیبی بدون انحناء

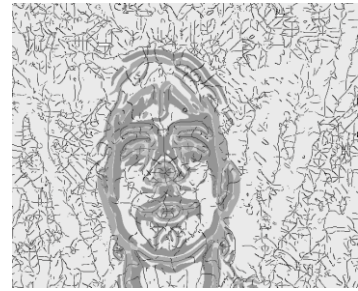
شکل ۳-۲۷(الف): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(الف)



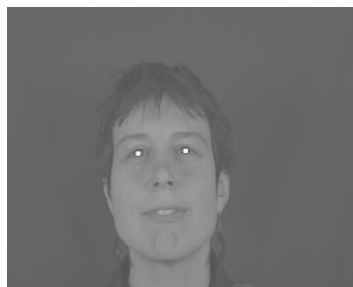
توزیع برجسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برجسب خورده توپوگرافیکی



پاسخ الگوریتم

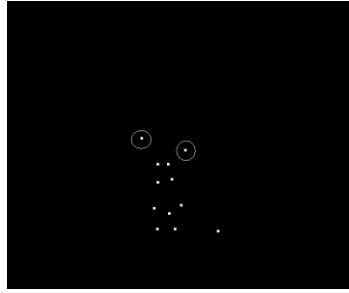


توزیع برجسب سراسیبی بدون انحناء

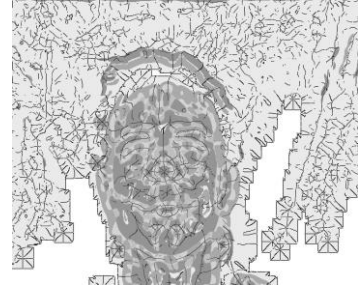
شکل ۳-۲۷(ب): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(ب)



توزیع برچسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برچسب خورده توپوگرافیکی



پاسخ الگوریتم

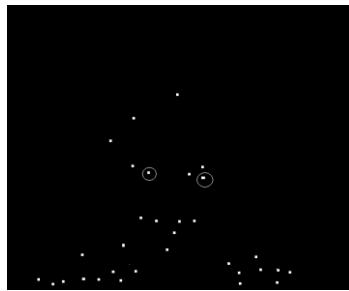


توزیع برچسب سراسیابی بدون انحنا

شکل ۳-۲۷(ج): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(ج)



توزیع برچسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برچسب خورده توپوگرافیکی



پاسخ الگوریتم

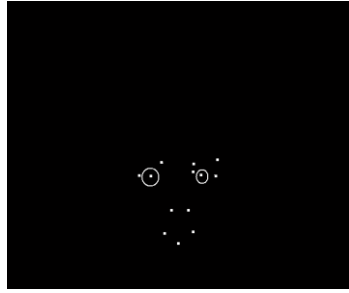


توزیع برچسب سراسیابی بدون انحنا

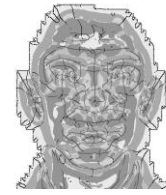
شکل ۳-۲۷(د): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(د)



توزیع برچسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برچسب خورده توپوگرافی

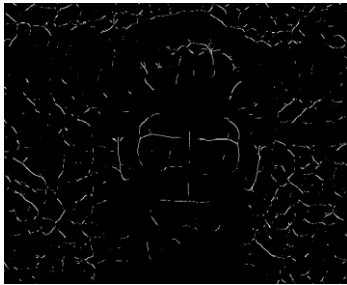


پاسخ الگوریتم

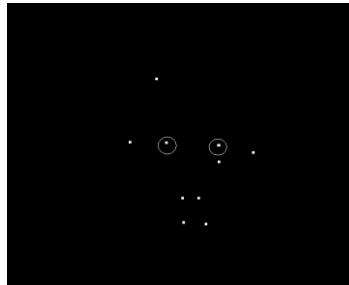


توزیع برچسب سرایشی بدون انحنا

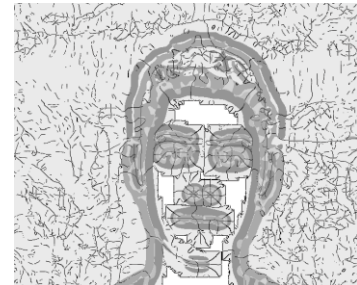
شکل ۳-۲۷(ه): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(ه)



توزیع برچسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برچسب خورده توپوگرافی



پاسخ الگوریتم

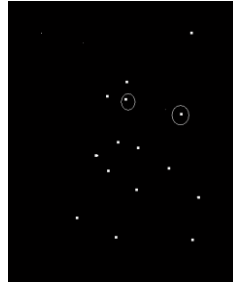


توزیع برچسب سرایشی بدون انحنا

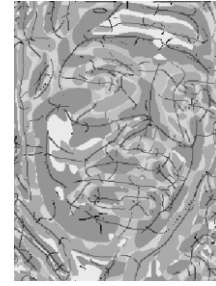
شکل ۳-۲۷(و): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(و)



توزیع برچسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برچسب خورده توپوگرافیکی



پاسخ الگوریتم

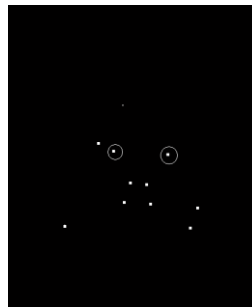


توزیع برچسب سراسیبی بدون انحنا

شکل ۳-۲۷(ی): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(ی)



توزیع برچسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برچسب خورده توپوگرافیکی

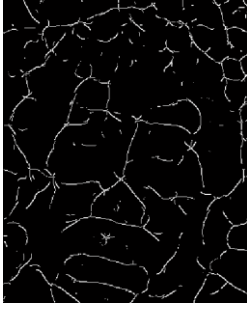


پاسخ الگوریتم

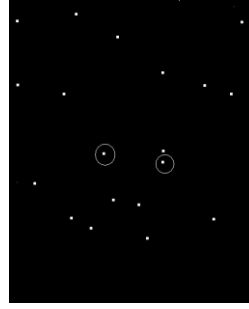


توزیع برچسب سراسیبی بدون انحنا

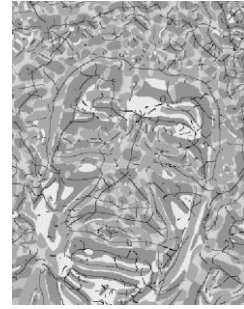
شکل ۳-۲۷(ن): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(ن)



توزیع برچسب ناودانی بیرون



نقاط کاندید مردمک



تصویر برچسب خورده توپوگرافیکی



پاسخ الگوریتم



توزیع برچسب سراسیابی بدون انحناء

شکل ۳-۲۷(م): نتایج مراحل مختلف الگوریتم پیشنهادی برای شکل ۲۱-۳(م)

شکل ۳-۲۷: نتایج الگوریتم پیشنهادی در شرایط مختلف

فصل ۴- ردیابی چشم در یک دنباله ویدیویی

در بسیاری از کاربردهای بینایی ماشین نظیر تعامل انسان و کامپیوتر، مطالعه خواب آلودگی راننده و... دنبال کردن چشم‌ها در فریم‌های متوالی یک رشته ویدئویی ضروری می‌باشد.

بدلیل کاربرد زیادی که سیستم‌های ردیاب شیء دارند تحقیقات زیادی در زمینه ردیابی الگو و اشیاء در پردازش تصویر انجام شده‌است. در ادامه مروری بر روشهای ردیابی^۱ مرتبط با الگوریتم ارائه شده در این تحقیق انجام می‌شود و سپس الگوریتم ردیابی بکار رفته با جزئیات بیشتری بررسی می‌شود.

۴-۱- مروری بر روشهای ردیابی

روش‌هایی که برای ردیابی مردمک مورد استفاده قرار می‌گیرند، به دلیل سادگی شکل این عضو و عاری بودنش از جزئیات پیچیده (برخلاف اشیائی مثل خودرو، شخص و غیره) در دسته روشهای مربوط به ردیابی اشیاء ساده قرار می‌گیرند. ردیابی این اشیاء معمولاً با محاسبه حرکت شیء بر اساس ناحیه در برگیرنده آن از هر فریم به فریم بعدی صورت می‌گیرد. حرکت شیء عموماً به فرم پارامتری (انتقال^۲، یکنوا^۳، نسبت صلبی^۴ و غیره) و یا توزیع یکنوا^۵ در فریم‌های پی‌اپی مدل می‌شود. این الگوریتم‌ها بر حسب ظاهر و تعداد اشیائی که می‌بایست ردیابی شوند و همچنین روشی که برای تخمین حرکت آنها انتخاب می‌گردد، متفاوت خواهند بود. این روشها در حالت کلی به دو دسته مدل‌های ظاهری مبتنی بر چگالی و الگو و مدل‌های ظاهری چند نمای^۵ تقسیم می‌شوند. بدلیل اینکه الگوریتم پیشنهادی در دسته اول قرار می‌گیرد، به ارائه توضیحاتی در رابطه با عملکرد روش‌های موجود در این دسته می‌پردازیم.

۴-۱-۱- ردیابی با استفاده از الگو و مدل‌های مبتنی بر چگالی

این روشها به دلیل سادگی بطور گسترده‌ای در ردیابی اشیاء مورد استفاده قرار می‌گیرند. بسته به اینکه این تکنیک‌ها بایستی اشیاء را به صورت جداگانه و یا مشترک ردیابی کنند به دو طبقه تک شیئی^۶ و چند شیئی^۷ تقسیم می‌شوند. در این تحقیق روش ردیابی تک شیئی مورد بررسی قرار می‌گیرد.

متداول‌ترین روش ردیابی اشیاء بصورت منفرد، تطبیق الگو^۸ است. در این تکنیک در تصویر فریم جدید، I_w دنبال ناحیه‌ای مشابه با الگوی شیء مورد نظر O_t ، که در فریم قبلی مشخص شده‌است

¹ Tracking

² Translation

³ Conformal

⁴ Affine

⁵ Multiview

⁶ Single Object

⁷ Multiple Objects

⁸ Template matching

می‌باشیم. مکان این الگو در تصویر کنونی با محاسبه معیاری برای تشخیص می‌زان شباهت صورت می‌گیرد. بعنوان مثال می‌توان همبستگی^۱ (رابطه ۴-۱) را بعنوان چنین معیاری در نظر گرفت. همبستگی متقابل بین دو سیگنال روشی جهت اندازه‌گیری شباهت و شناسایی الگوها می‌باشد [۵۴]. در این رابطه (dx, dy) جابجائی مکان الگو در فریم جدید را نشان می‌دهند.

$$\arg \max_{dx, dy} \frac{\sum_x \sum_y (O_t(x, y) \times I_w(x + dx, y + dy))}{\sqrt{\sum_x \sum_y O_t^2(x, y)}} \quad 1-4$$

معمولاً شدت روشنایی تصویر و ویژی‌های رنگ برای تعریف الگو استفاده می‌گردد. به دلیلی اینکه شدت روشنایی تصویر بسیار حساس به تغییرات نور محیط است از گرادیان روشنایی تصویر نیز می‌توان استفاده نمود زیرا حساسیت کمتری به تغییرات نور محیط دارد.

یکی از محدودیت‌های تطبیق الگو، حجم محاسبات بالا در آن است که به دلیل نیاز به جستجوی دقیق در تصویر، ایجاد می‌شود. برای کاهش حجم محاسبات معمولاً ناحیه جستجوی الگوریتم در فریم جدید محدود به همسایگی شیء در فریم قبلی می‌گردد نه کل تصویر.

همانطور که گفته شد استفاده از معیار همبستگی برای ردیابی محاسبات زیادی در پی دارد. در بسیاری از حوزه‌ها به کمک تحلیل در حوزه فرکانس (تبدیل فوری) سعی در سرعت دادن به محاسبات دارند. اما متأسفانه فرم نرمالیزه شده همبستگی (ضرایب همبستگی) که برای اهداف تطبیق الگو مناسب است معادل ساده و کارامدی در حوزه فرکانس ندارد. به این دلیل همبستگی متقابل نرمالیزه شده در حوزه ابعاد تصویر محاسبه می‌گردد [۵۴]. به دلیل محاسبات پیچیده لازم برای بدست آوردن همبستگی در حوزه ابعاد، برخی روشهای تقریبی ولی سریعی در این حوزه برای اهداف تطبیق الگو ارائه شده است [۵۵]. Lewis [۵۶]، روشی را برای تطبیق الگو با معیار فاصله زیر پیشنهاد می‌کند:

$$d_{ft}^2 = \sum_{x,y} [f(x, y) - t(x - u, y - v)]^2 \quad 2-4$$

در این رابطه f تصویر ورودی و t الگویی است که در تصویر دنبال آن می‌گردیم. الگوی t در محل u و v قرار گرفته است. با گسترش رابطه خواهیم داشت:

$$d_{f,t}^2(u, v) = \sum_{x,y} [f^2(x, y) - 2f(x, y)t(x - u, y - v) + t^2(x - u, y - v)] \quad 3-4$$

ترم $\sum_{x,y} t^2(x - u, y - v)$ در عبارت بالا برابر با انرژی الگو می‌باشد و ثابت است.

در صورتی که انرژی تصویر $\sum_{x,y} f^2(x, y)$ تقریباً ثابت و مستقل از v, u باشد، عبارت باقی‌مانده زیر معیاری برای اندازه‌گیری شباهت بین تصویر و الگو در محل v, u می‌باشد.

¹ Correlation

$$c(u,v) = \sum_{x,y} f(x,y)t(x-u,y-v) \quad 4-4$$

استفاده از معیار شباهت برای تطبیق الگو چندی نضعف دارد:

۱. اگر انرژی تصویری $\sum_{x,y} f^2(x,y)$ نسبت به مکان تغییری کند، تطبیق با استفاده از رابطه

فوق دچار خطا خواهد شد. برای مثال همبستگی بین الگو و یک ناحیه منطبق با الگو در

تصویری ممکن است کمتر از همبستگی بین الگو و یک ناحیه روشن تصویر باشد.

۲. محدوده تغییرات معیار شباهت $c(u,v)$ بستگی به ابعاد الگو دارد و کمی متغیر نیست.

۳. معیار شباهت فوق نسبت به تغییرات روشنایی تصویری در دو فریم متوالی پایدار نیست.

با نرمالیزه کردن تصویری و الگو از لحاظ روشنایی شباهت به کمک همبستگی بصورت زیر تعریف

میگردد، این معیار مشکلات معیار قبلی را ندارد.

$$\gamma(u,v) = \frac{\sum_{x,y} [f(x,y) - \bar{f}_{u,v}][t(x-u,y-v) - \bar{t}]}{\left\{ \sum_{x,y} [f(x,y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x-u,y-v) - \bar{t}]^2 \right\}^{0.5}} \quad 5-4$$

در این رابطه \bar{t} میانگین الگو و $\bar{f}_{u,v}$ میانگین $f(x,y)$ در محل قرار گرفتن الگو (u,v) است. به

رابطه ۵-۴ تابع همبستگی متقابل نرمالیزه شده^۱ گفته می شود.

در روش دیگری که توسط Adachi و همکارش [۵۷] پیشنهاد شد، فرایند ردیابی در دو مرحله

بصورت زیر انجام می شود:

۱. پیدا کردن پنجره ای که بیشترین همبستگی را با الگو دارد. این کار با محاسبه خطای

تطبیق (حاصل جمع قدر مطلق تفاضل^۲ (SAD)) صورت می گیرد.

$$Dist[i,j] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} |R[k,l] - M[i+k,j+l]|, \quad i,j: 0 \leq i,j \leq 15, \quad 6-4$$

که در این رابطه $R[x,y]$ و $M[x,y]$ به ترتیب الگو و پنجره ای از تصویر است که با الگو داده می شود.

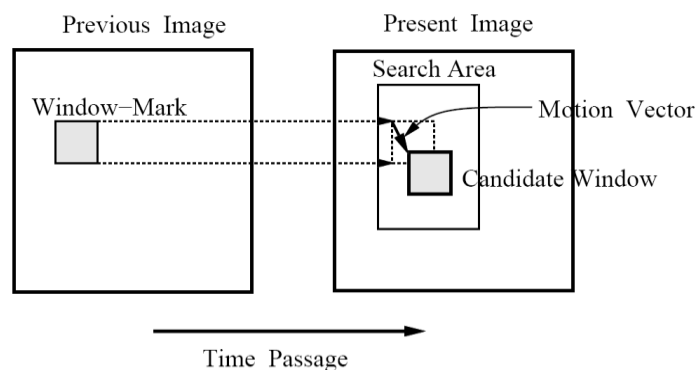
L و K به ترتیب مربوط به ارتفاع و عرض پنجره الگو می باشند.

۲. حرکت دادن پنجره به مکان جدید.

شکل ۴-۱ این روند را به خوبی نشان می دهد.

¹ Normalized Cross-Correlation Function (NCCF)

² Sum of Absolute Difference (SAD)



شکل ۱-۴: ردیابی در روش Adachi و همکارش [۵۷]

یکی از الگوریتم‌های پی‌چیده بر مبنای تطبیق الگو توسط Schweitzer و همکارانش [۵۸] ارائه گردید. در این گونه روشها علاوه بر سطح روشنایی الگو از سایه و ویژگی‌ها نظیر فراوانی رنگ و یا مدل‌های مخلوط به عنوان ویژگی الگو در همسایگی بی‌ضوی و یا مستطیلی تصویر انتخاب می‌گردند.

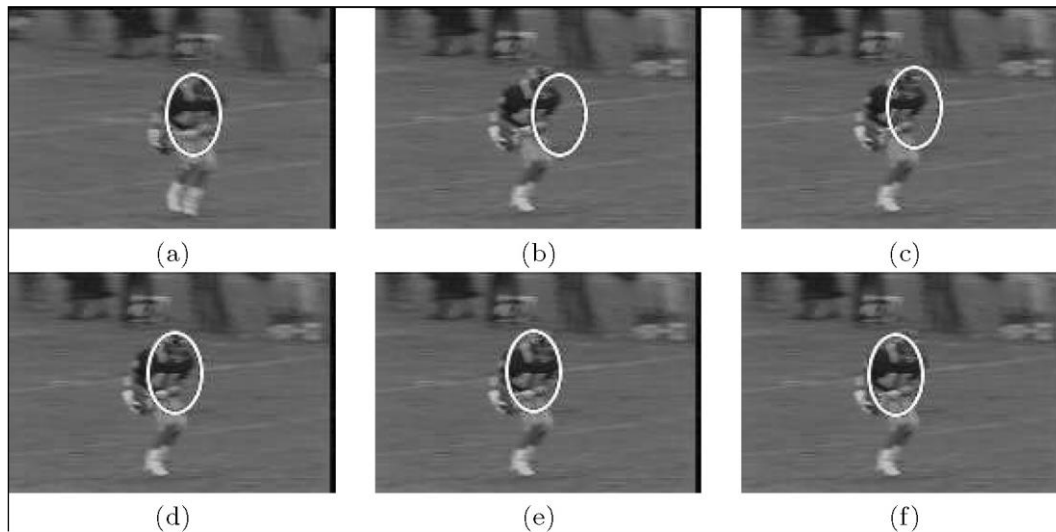
Fieguth و همکارش [۵۹] با پیدا کردن می‌انگین رنگ پی‌کسل‌های واقع در ناحیه مربعی مربوط به الگو، مدل الگو را تعریف کردند. آنها برای کاهش پی‌چیدگی محاسباتی الگو را در هشت ناحیه همسایه مورد بررسی قرار دادند. در روش آنها شباهت بین الگو و پنجره‌ای فرضی در تصویر با محاسبه نسبت بین می‌انگین رنگها در این دو ناحیه تعیین می‌شود. مکانی از تصویر که بیشترین شباهت را ایجاد کند به عنوان محل صحیح الگو در تصویر انتخاب می‌شود.

Comanicu و همکارش [۶۰] از فراوانی وزن یافته که از روی ناحیه دایره‌ای شکلی در اطراف شیء محاسبه می‌شود به عنوان الگو استفاده کردند. سپس به جای انجام جستجوی کلی در تصویر برای مکانیابی الگو از روش جابجایی میانگین^۱ استفاده نمودند. ردیاب جابجایی میانگین در فرایندی تکراری با مقایسه هیستوگرام الگو و پنجره‌ای فرضی در تصویر، مکانی را که در آن شباهت ظاهری بیشینه می‌گردد بدست می‌آورد. شباهت بین هیستوگرام از روی ضرایب $Bhattacharya$ ، $\sum_{u=1}^b P(u)Q(u)$ ، (b تعداد ستونهای هیستوگرام) تعیین می‌شود. در هر تکرار، بردار جابجایی میانگین طوری تعریف می‌گردد که شباهت هیستوگرام افزایش یابد. این فرایند تا همگرا شدن الگوریتم ادامه پیدا می‌کند که معمولاً بعد از حدود ۵ تا ۶ تکرار خواهد بود.

Comanicu و همکارش [۶۰] برای تولید هیستوگرام از روش وزن دهی توسط یک کرنل در حوزه فاصله استفاده می‌کنند. این کرنل به پی‌کسل‌های که به الگو نزدیکترند وزن بیشتری اختصاص می‌دهد. Comanicu [۶۱] روش ردیابی براساس جابجایی میانگین را با استفاده از هیستوگرام رنگ-فاصله [۶۰] به جای هیستوگرام رنگ به تنهایی، توسعه داد. ویژگی بارز استفاده از جابجایی میانگین نسبت به روشهای مبتنی بر الگوی استاندارد، عدم جستجوی وسیع و وقتگیری در تصویر و همچنین یافتن مکان الگو در تعداد اندکی تکرار می‌باشد. با این حال شرط لازم برای عملکرد مطلوب جابجایی الگو این است که در اولین تکرار حداقل بخشی از الگوی مورد نظر

¹ Mean Shift Algorithm

می بایست داخل ناحیه دایره‌ای جستجوی انتخاب شده قرار گیرد. مثالی از ردیابی الگو به کمک جابجایی میانگین در شکل ۴-۲ آمده است.



شکل ۴-۲: تصویرهای مربوط به تکرارها در روش جابجایی میانگین (a) مکان تخمین شیء در مرحله $t-1$ (b) فریم در مرحله t با توجه به محل اولی که در تکرار قبلی تعیین شد. (f) محل نهایی شیء بعد از ۶ تکرار

Jepson و همکارانش [۶۲] روشی برای ردیابی شیء، براساس ترکیبی از سه جزء شامل ویژگی‌های ظاهری پایدار، ویژگی‌های گذرا و فرایندی نویزی ارائه می‌دهند. جزء پایدار، قابل اعتمادترین ویژگی (ظاهر) را برای تخمین حرکت تشخیص می‌دهد که در واقع نواحی از الگو می‌باشند که ظاهر آنها در فریم‌های متوالی تغییری چندانی نمی‌کنند. جزء گذرا پی‌کسل‌های با تغییری روشنایی ناگهانی را تشخیص می‌دهند. جزء نویزی تشخیص نواحی خارجی اطراف الگو را که نویز مانند می‌باشد بر عهده دارد. Jepson و همکارانش از الگوریتم EM^۱ برای یادگیری پارامترهای این سه جزء استفاده می‌کنند. آنها از فاز پاسخ فیلتر قابل چرخش^۲ به عنوان ویژگی برای توصیف ظاهر استفاده می‌کنند. شکل شیء به صورت یک بیضی روی تصویر نمایش داده می‌شود و حرکت شیء از روی پی‌چش^۳ ناحیه ردیابی شده از هر فریم به فریم بعدی محاسبه می‌گردد [۶۰].

۴-۱-۲- الگوریتم تطبیق بلوک (BMA)

این الگوریتم شامل تفکیک هر فریم به بلوک‌هایی با ابعاد ثابت (۶×۶ یا ۸×۸ یا ...) و شناسایی جابجایی بلوکها بین فریم کنونی و فریم پیشین است. این روش می‌دانی از بردارهای جابجایی (DVF)^۵ که جهت حرکت شیء را نشان می‌دهد ایجاد می‌کند. در این الگوریتم هر بلوک بخشی از تصویر را در بر می‌گیرد و بصورت یک ماتریس که شامل ضرایب سطح خاکستری این بخش از

¹ Expectation Maximization

² Steerable Filter

³ Warping

⁴ Block matching algorithm

⁵ Field of Displacement Vectors

تصویری است بی‌ان می‌شود. کاری که باید انجام شود تخمین موقعیت هر بلوک در فریم قبلی به منظور محاسبه جابجایی و سرعت (Δt بین فریمها) است. هر بلوک، فضای جستجو را در فریم قبلی نشان می‌دهد که مرکز آن با مرکز بلوک منطبق است. سپس بلوک پی‌کسل به پی‌کسل در داخل ناحیه جستجو شیفت پیدا می‌کند. به ازای هر شیفت می‌زان تطابق محاسبه می‌گردد. هدف از این مقایسه تعیین مجموعه پی‌کسل‌های است که بیشترین شباهت را در ناحیه جستجو با بلوک فریم قبلی دارد. با توجه به مطالب عنوان شده، در اینجا روش تابع همبستگی متقابل هنجار شده انتخاب می‌گردد.

$$NCCF = \frac{\sum_{i,j} [P(i,j) \cdot q(i,j)]}{\sqrt{\left[\sum_{i,j} P^2(i,j) \right] \cdot \left[\sum_{i,j} q^2(i,j) \right]}} \quad 7-4$$

در رابطه 7-4، "بی‌انگر حاصلضرب یک به یک عناصر ماتریسهای P و q می‌باشد. ابعاد ناحیه جستجو باید طوری تعیین گردد که بیشترین جابجایی ممکن برای الگو را در دو فریم متوالی پوشش دهد. البته به دلیل زمان محاسباتی این ناحیه باید تا حد امکان کوچک انتخاب گردد. فرایند تطبیق با استفاده از محاسبه همبستگی، بردار جابجایی V' بین مکان بلوک در فریم جاری و نزدیکی نقطه به آن را در فریم قبلی ایجاد می‌کند. در فریم جاری قرینه این بردار یعنی $V' = -V$ که در مرکز بلوک قرار گرفته است جابجایی را به همراه اطلاعات ردیابی نشان می‌دهد. ماتریس DVF شامل یک بردار به ازای هر مکان است. بنابراین پس از اعمال BMA هر بلوک به طور جداگانه ردیابی شده است. الگوریتم برچسب زنی اطلاعات لازم برای ردیابی را مشخص می‌کند. این اطلاعات ناحیه‌ای بزرگتر از همان بلوکهای مرحله قبل هستند که ویژگیهای جابجایی مشابه دارند. به بیانی دقیق‌تر الگوریتم برچسب زنی لیست بلوکهایی را که متعلق به یک شی هستند مشخص می‌کند و از روی خروجی BMA می‌توان تشخیص داد که هر بلوک از چه مکانی (با توجه به فریم قبلی) به مکان فعلی منتقل شده است. با الحاق این بلوکها می‌توان کل شیء را ردیابی کرد.

متأسفانه در شرایطی که تصویری نویزی باشد روش BMA دچار خطا می‌گردد. برای کاهش این اثر می‌بایست سهم بردارهایی که تطبیق مناسبی نداشته‌اند را در تعیین نهایی جهت حرکت حذف کرد [63]. برای اینکار معمولاً مقدار ثابتی از پارامترهای بدست آمده با تطبیق ضعیف کسر می‌گردد. با این کار برای بلوکهایی که تطبیق خوبی در ناحیه جستجو ندارند مقدار جابجایی صفر در نظر گرفته خواهد شد. در ادامه ابتدا همبستگی بلوک جاری با بلوکی هم اندازه و هم مکان در فریم قبلی محاسبه شده (یعنی جابجایی صفر) و سپس با یک آستانه ثابت مقایسه می‌گردد. اگر همبستگی بیشتر از آستانه باشد بردار جابجایی صفر برای این بلوک در نظر گرفته می‌شود و در غیر این صورت محاسبه این همبستگی برای سایر بلوکهای ممکن در محدوده جستجو انجام گرفته و ماکزیمم آن انتخاب می‌گردد. پس از اعمال آستانه در صورتی که شرط بیشتر بودن همبستگی از آستانه برقرار باشد بردار جابجایی برابر با اختلاف مکان قبلی و مکان فعلی بلوک خواهد بود.

در روشی که توسط Stefano و همکارانش [۶۴] ارائه شد، روند ردیابی براساس الگوریتم تطبیق بلوک^۱ (BMA) انجام می شود.

در بخش بعد الگوریتم ارائه شده در این تحقیق برای ردیابی چشمها مورد بررسی قرار می گیرد.

۴-۲- الگوریتم پیشنهادی برای ردیابی چشمها در دنباله تصاویر

الگوریتم ارائه شده برای ردیابی چشمها از نوع تطبیق الگو می باشد که ناحیه تعیین شده بعنوان چشمها در مرحله تشخیص چشم را به عنوان الگو در نظر گرفتیم .

مراحل مختلف این الگوریتم در بلوک دیاگرام شکل ۳-۴ نشان داده شده است.

در اولین فریم ورودی به کمک الگوریتم تشخیص چشم محل چشمها استخراج شده و عملیات ردیابی با استفاده از مکان چشمها در اولین فریم آغاز می گردد.

در این الگوریتم، عملیات تطبیق الگو بر پایه اطلاعات سطح خاکستری تصویر انجام می شود. به این ترتیب لزوم تعیین ساختار توپوگرافیکی در همه فریمها از بین می رود. این موجب سرعت بخشیدن به عمل ردیابی چشمها می شود.

مراحل ردیابی مکان چشمها به ترتیب زیر می باشند:

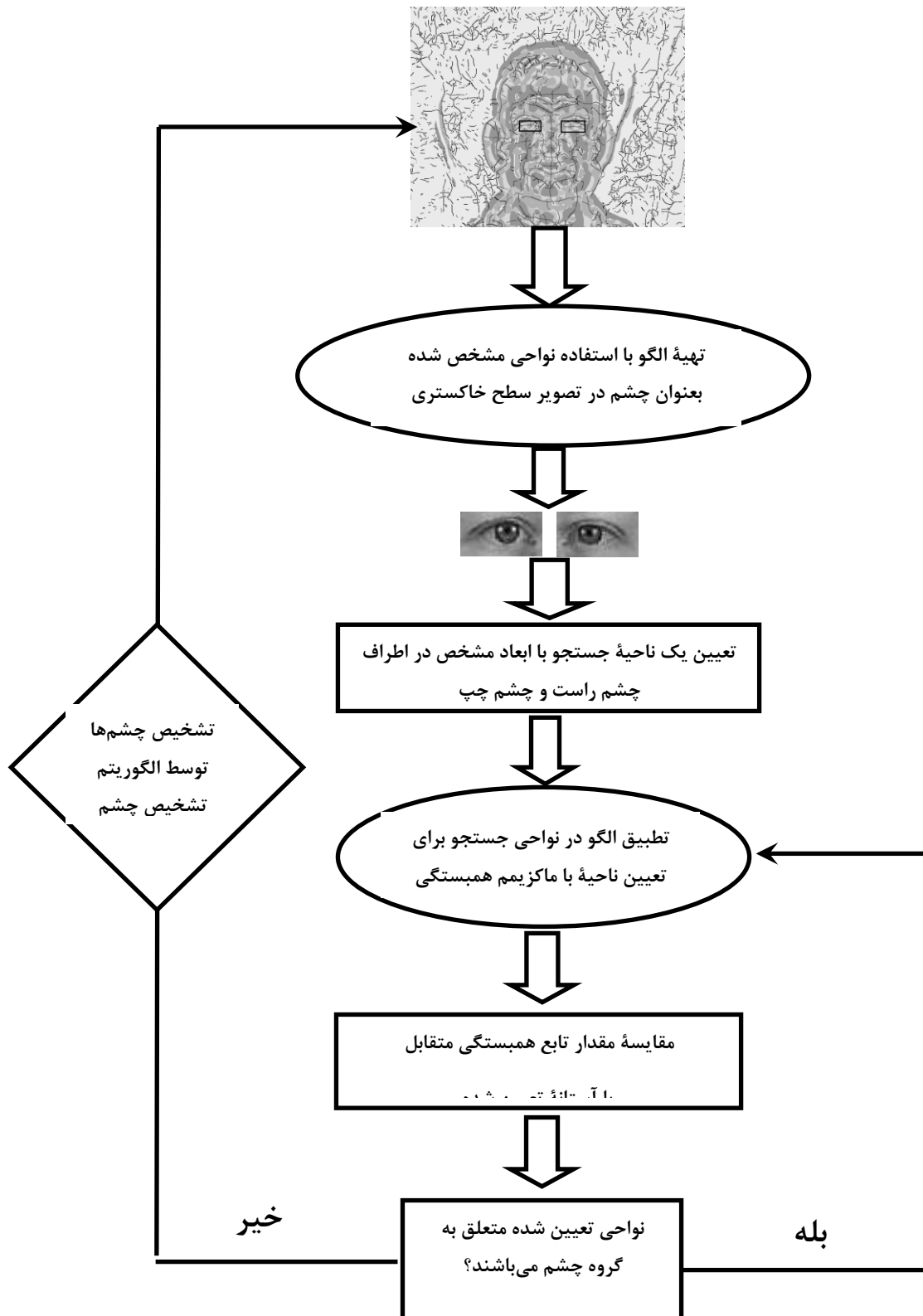
۱. پس از تعیین مکان چشمها در اولین فریم، ناحیه تعیین شده بعنوان چشم راست و چشم چپ از تصویر سطح خاکستری جدا شده و بعنوان الگو برای تشخیص چشمها در فریمهای بعدی مورد استفاده قرار می گیرند.

۲. تطبیق الگو در فریم جدید به منظور تعیین ناحیه چشمها انجام می گیرد. همانطور که در بخش قبل گفته شد برای اجتناب از محاسبات زمان بر، ناحیه جستجو به بخش کوچکی در اطراف چشمها در فریم قبل محدود می شود. در الگوریتم پیشنهادی ابعاد ناحیه جستجو $0.5d \times 0.5d$ در نظر گرفته شد که d فاصله بین دو مردمک می باشد. دلیل استفاده از ناحیه جستجو، با دقت به این نکته صورت می گیرد که سر در دو فریم متوالی به میزان زیادی جابجا نمی شود. برای تعیین میزان شباهت بین الگو و تصویر در ناحیه جستجو از تابع همبستگی^۲ نرمالیزه شده استفاده می شود و ناحیه ای که مقدار این تابع در آن ماکزیمم ممکن مقدار را داشته باشد بعنوان چشم در فریم جدید انتخاب می شود.

¹ Block Matching Algorithm (BMA)

² Correlation

۳. برای تأیید نواحی انتخاب شده بعنوان چشم از مقدار بدست آمده توسط تابع همبستگی متقابل استفاده می‌گردد. در صورتیکه این مقدار از یک آستانه معین کوچکتر باشد نشان دهنده خطا در نتیجه می‌باشد و بایستی مکان چشم‌ها مجدداً توسط الگوریتم تشخیص چشم تعیین شود. مقدار آستانه در الگوریتم پیشنهادی 0.65 در نظر گرفته شده است.

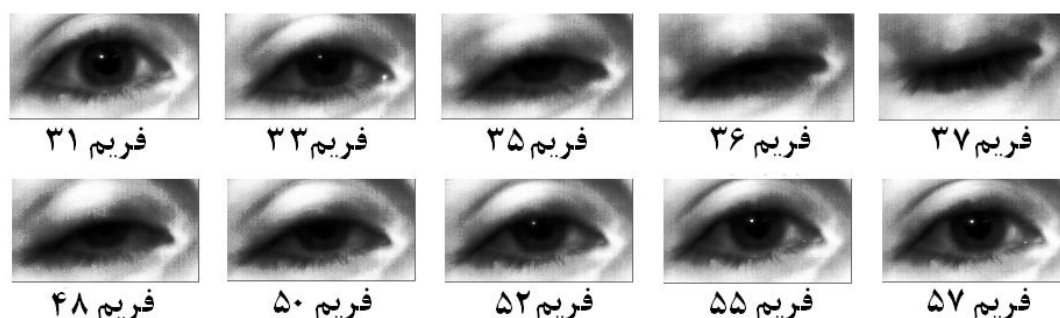


شکل ۳-۴: بلوک دیاگرام الگوریتم پیشنهادی برای ردیابی چشم‌ها

فصل ۵- تعیین فاصله بین پلک‌ها در چشم

با تشخیص محل مردمک و استخراج ناحیه چشم‌ها، بررسی حرکت اجزای چشم امکان‌پذیر می‌شود. پلک‌ها اجزایی از چشم هستند که تحلیل حرکت آنها نسبت به هم حاوی اطلاعات مفیدی در رابطه با رفتار شخص می‌باشد.

پلک زدن یکی از اعمال طبیعی چشم‌ها است که برای مدتی کوتاه موجب پوشش سطح چشم‌ها می‌شود. در این فرایند پلک بالا سهم بیشتری دارد شکل ۵-۱، تعدادی از تصاویر یک دنباله ویدیویی مربوط به فرایند پلک زدن را نشان می‌دهد.



شکل ۵-۱: تصاویر تغییر حالت چشم در پلک زدن

یکی دیگر از اعمال پلک‌ها که مشابه با پلک زدن می‌باشد، حرکت سریع یا لرزش پلک^۱ است. در مقایسه با دو رفتار ذکر شده، حالت دیگری نیز از نحوه حرکت پلک‌ها وجود دارد که مهمترین فاکتور در تشخیص خواب می‌باشد، منظور از این حالت بسته شدن کامل پلک‌ها می‌باشد.

متمایز نمودن پلک زدن از بسته شدن کامل پلک‌ها اهمیت زیادی دارد زیرا هرکدام از آنها مفهوم جداگانه‌ای دارد. با مطالعه حرکت پلک‌ها در طول زمان جدا سازی این دو رفتار فراهم می‌شود. پلک زدن، پنهان شدن گذرای عنبیه را در پی دارد در صورتیکه مدت زمان ناپدید شدن عنبیه در تصاویر مربوط به خواب و بسته شدن کامل پلک‌ها در مقایسه با پلک زدن بسیار بیشتر می‌باشد. بر اساس تحقیقات انجام شده توسط Thorslund [۶۵] فرکانس پلک زدن متغیر است و تحت تأثیر عواملی از قبیل وضع روانی^۲، خستگی و ... می‌باشد. در حالتی که شخص استرس نداشته باشد، نرخ پلک زدن ۲۰-۱۵ بار در هر دقیقه می‌باشد و در زمان مطالعه این عدد به ۳ بار در دقیقه کاهش می‌یابد.

تشخیص پلک زدن قدم مهمی در توسعه سیستم‌های مربوط به تعامل انسان و کامپیوتر، بخصوص سیستم‌های تسهیل کننده ارتباط افراد معلول و ناتوان با کامپیوتر می‌باشد. در این سیستم‌ها از پلک زدن برای کنترل مکان‌نما استفاده می‌شود. چنین سیستم‌هایی باید قادر به ایجاد مکانی‌زمی برای جلی‌گزینی کردن دکمه‌های ماوس باشد.

¹ Eyelid Fluttering

² Mood State

در کاربردهایی از قبیل تشخیص خواب‌آلودگی رانندگان یا تشخیص بیماری‌ها، تشخیص بسته شدن پلک‌ها به مدت طولانی (در مقایسه با پلک زدن) و جدا کردن آن از پلک زدن شخص اهمیت بسیار زیادی دارد. با توجه به اینکه تشخیص خواب در رانندگان در محیط ماشین و در شرایط گوناگونی از نظر نورپردازی اتفاق می‌افتد، لذا روش‌های ارائه شده در این حوزه باید مقاومت زیادی نسبت به تغییر شرایط نورپردازی داشته باشند.

در پیاده‌سازی الگوریتم‌های تشخیص پلک زدن یا تشخیص خواب‌آلودگی، اولین قدم پس از تعیین ناحیه چشم‌ها مشخص نمودن وضعیت پلک‌ها می‌باشد. با ردیابی ناحیه چشم‌ها در طول زمان، استخراج پلک‌ها و تعیین فاصله بین آنها امکان‌پذیر می‌باشد.

در روش پیشنهادی با استفاده از اطلاعات مربوط به تغییر سطح روشنایی در ناحیه چشم به استخراج محل پلک‌ها در تصویر می‌پردازیم. شکل ۲-۵ تصویر سطح خاکستری چشم را نشان می‌دهد. همانطور که در شکل نیز مشاهده می‌شود، سطح روشنایی پیکسل‌ها در ناحیه چشم به میزان قابل توجهی پایین‌تر از سطح روشنایی پوست می‌باشد و پلک‌ها مرز تغییر سطح روشنایی بین ناحیه چشم و پوست می‌باشند.



شکل ۲-۵: تصویر چشم در سطح خاکستری

۱-۵- توابع نگاشت^۱

بطور کلی، توابع نگاشت ابزارهای مفیدی برای آشکارسازی مرز نواحی مختلف تصویر می‌باشند. فرض کنید PF یک تابع نگاشت و ξ یک ثابت عددی کوچک باشد. اگر مقدار PF بطور ناگهانی از z_0 به $z_0 + \xi$ تغییر کند، آنگاه z_0 می‌تواند یک مرز بین دو ناحیه همگن باشد.

بعنوان مثال مرزهای عمودی در یک تصویر با استفاده از تابع نگاشت در جهت عمودی و یک آستانه بصورت زیر تعیین می‌شوند:

$$\theta_v = \max \left\{ \left| \frac{\partial PF_v(x)}{\partial x} \right| > T \right\} \quad 1-5$$

اندیس v برای تابع PF، نشان می‌دهد که تابع نگاشت در راستای محور y (جهت عمودی) تصویر محاسبه شده‌است. کاملاً بدیهی است که با محاسبه تابع نگاشت در راستای محور x (جهت افقی) تصویر مرزهای افقی بطور مشابه با مرزهای عمودی تعیین می‌شوند.

¹ Projection Function

۵-۱-۱- تابع نگاشت جمع کننده

تابع نگاشت جمع کننده^۱، یکی از توابع نگاشت می‌باشد. تابع نگاشت جمع کننده در جهت عمودی و افقی به ترتیب با $IPF_v(x)$ و $IPF_h(y)$ نشان داده می‌شود و با روابط زیر محاسبه می‌شود:

$$IPF_v(x) = \sum_{y_i=y_1}^{y_2} I(x, y_i) \quad ۲-۵$$

$$IPF_h(y) = \sum_{x_i=x_1}^{x_2} I(x_i, y) \quad ۳-۵$$

در این روابط، $I(x,y)$ شدت روشنایی پیکسل در مکان (x,y) تصویر می‌باشد. در روابط ۲-۵ و ۳-۵ محدوده محاسبه تابع نگاشت جمع کننده در جهت عمودی بازه $[y_1, y_2]$ و محدوده محاسبه تابع نگاشت جمع کننده در جهت افقی بازه $[x_1, x_2]$ می‌باشد.

تابع نگاشت جمع کننده (IPF)، متداول‌ترین تابع نگاشت مورد استفاده می‌باشد.

۵-۱-۲- تابع نگاشت واریانس

تابع نگاشت واریانس^۲ (VPF) توسط Feng و همکارش [۶] پیشنهاد شد. تابع نگاشت واریانس نیز در دو جهت عمودی و افقی محاسبه می‌شود (روابط ۴-۵ و ۵-۵).

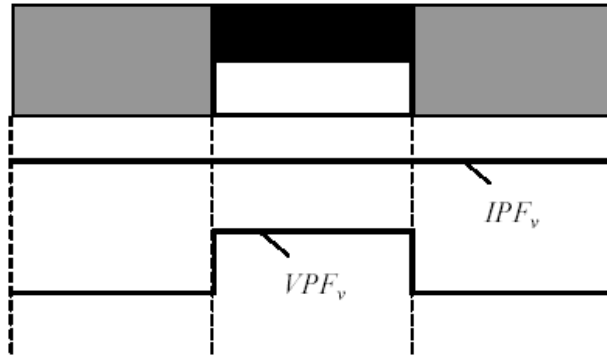
$$VPF_v(x) = \frac{1}{y_2 - y_1} \sum_{y_i=y_1}^{y_2} [I(x, y_i) - IPF'_v(x)] \quad ۴-۵$$

$$VPF_h(y) = \frac{1}{x_2 - x_1} \sum_{x_i=x_1}^{x_2} [I(x_i, y) - IPF'_h(y)] \quad ۵-۵$$

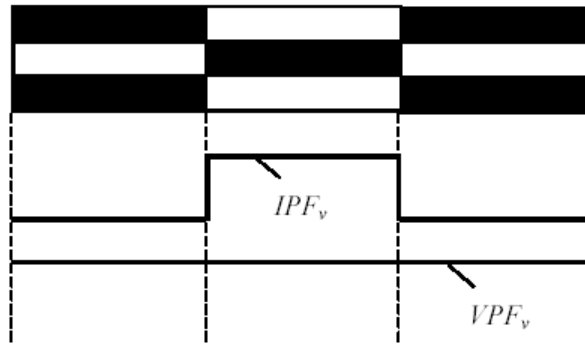
تابع نگاشت واریانس نسبت تابع نگاشت جمع کننده حساسیت بیشتری در مقابل وجود تغییرات در تصویر نشان می‌دهد [۶]، اما مواردی تابع نگاشت واریانس قادر به آشکارسازی مرز نواحی نمی‌باشد. شکل‌های ۳-۵ و ۴-۵ مثال‌هایی از نتایج بدست آمده توسط تابع نگاشت واریانس و تابع نگاشت جمع کننده را نشان می‌دهد. در شکل ۳-۵، استفاده از تابع نگاشت جمع کننده منجر به نتیجه خطا شده‌است در صورتیکه تابع نگاشت واریانس مرز بین نواحی مختلف را بخوبی آشکار کرده‌است. شکل ۴-۵ نیز مربوط به موردی است که استفاده از تابع نگاشت جمع کننده نتیجه بهتری در پی داشته است.

¹ Integral Projection Function

² Variance Projection Function



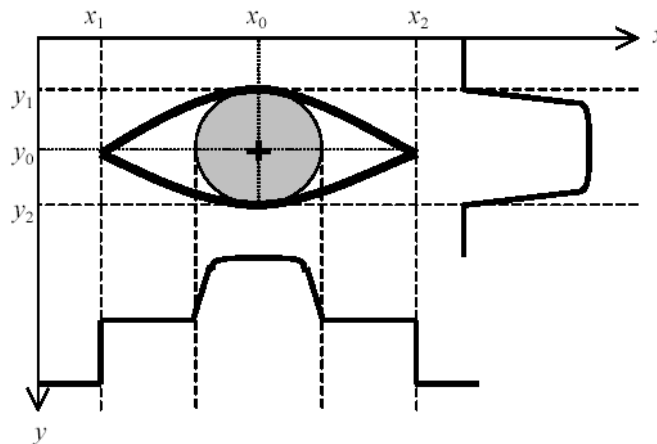
شکل ۳-۵: عملکرد مناسب VPF_v در تعیین مرزهای شکل



شکل ۴-۵: عملکرد مناسب IPF_v در تعیین مرزهای شکل

۵-۱-۳- تشخیص نواحی مختلف چشم توسط توابع نگاشت

مدل یک چشم در شکل ۵-۵ نشان داده شده است. مرزهای چشم در راستای عمودی و افقی به ترتیب توسط تابع نگاشت جمع کننده در جهت افقی $IPF_h(y)$ و تابع نگاشت جمع کننده در جهت عمودی $IPF_v(x)$ تعیین شده اند. همانطور که در شکل نیز مشخص است، مقدار این توابع در مرزهای چشم بطور ناگهانی تغییر می کند. مرزهای چشم در راستای محور x با x_1 و x_2 در شکل مشخص شده اند. این نقاط مؤلفه افقی مختصات گوشه های چشم می باشند. مرزهای چشم در راستای محور y نیز با y_1 و y_2 مشخص شده اند که این نقاط مؤلفه عمودی مختصات پلک ها می باشند.



شکل ۵-۵: مدل چشم

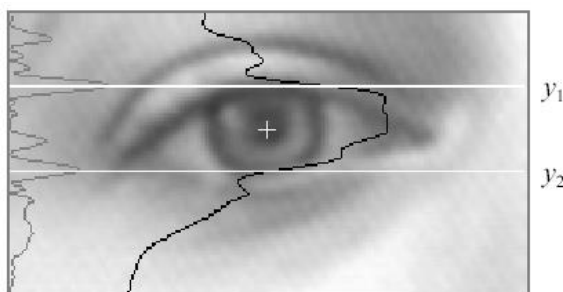
با مشخص شدن x_1, x_2, y_1 و y_2 مکان نقطه مرکزی چشم (x_0, y_0) بصورت زیر تعیین می‌شود:

$$x_0 = \frac{x_1 + x_2}{2}$$

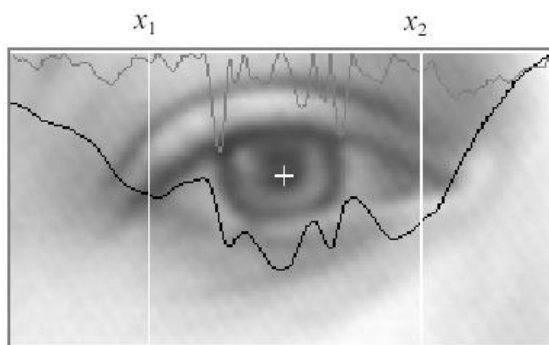
۶-۵

$$y_0 = \frac{y_1 + y_2}{2}$$

مثالی از تشخیص مرزهای چشم توسط تابع نگاشت جمع کننده در شکل ۶-۵ نشان داده شده است. منحنی مشکی، مقادیر محاسبه شده برای این تابع در نقاط مختلف تصویر را نشان می‌دهد و منحنی خاکستری مشتق مرتبه اول تابع نگاشت جمع کننده را می‌باشد.



(الف): تعیین مرزهای افقی



(ب): تعیین مرزهای عمودی

شکل ۶-۵: تعیین مرزهای چشم توسط تابع نگاشت جمع کننده

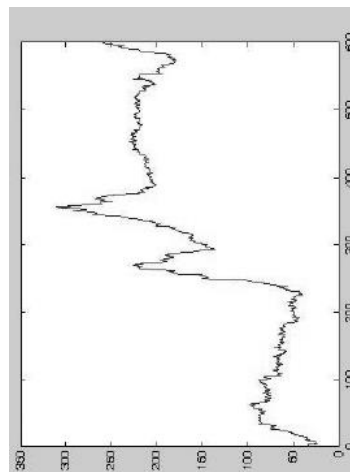
با توجه به اینکه ناحیه چشم تیره‌تر از نواحی همسایه‌اش می‌باشد، تابع نگاشت جمع کننده ابزار مناسبی برای استخراج مرز جدا کننده چشم از نواحی همسایه‌اش می‌باشد. این مرز جدا کننده، پلک‌ها هستند. مزیت مهم دیگر این روش پی‌چیدگی بسیار کم و سرعت بالای آن است. هر چند که سالی‌ها روشها از دقت بالاتری برخوردارند ولی پی‌چیدگی محاسباتی بالایی دارند. همچنین بعضی از روش‌ها که از تصویر برداری مادون قرمز استفاده می‌کنند برای پیاده‌سازی نیاز به تجهی‌زات مختلفی دارند.

Gan و همکارانش [۵] با استفاده از تابع نگاشت جمع کننده، روشی برای تشخیص خواب آلودگی در رانندگان ارائه دادند. در روش آنها در ابتدا ناحیه چهره از تصویر جدا شده و سپس تابع نگاشت جمع کننده در جهت افقی بر روی تصویر باینری چهره محاسبه می‌گردد. برای محاسبه این تابع پیکسل‌های با مقدار صفر شمارش می‌شوند. شکل ۷-۵ تصویر چهره را به همراه منحنی تابع نگاشت محاسبه شده نشان می‌دهد. محل چشمها در این منحنی نقاط با حداکثر دامنه می‌باشند و مرزهای افقی چشمها نیز (پلک بالا و پلک پایین) در دره‌های این منحنی واقع شده‌اند. در روش آنها محل پلک‌ها با استفاده از شیب منحنی تابع نگاشت تعیین شد. شیب در هر نقطه از این منحنی به صورت زیر محاسبه می‌شود:

اگر مختصات عمودی یک پیکسل y باشد و مقدار تابع نگاشت جمع کننده در این نقطه با $f(y)$ مشخص شود. مقدار شیب در نقطه y بصورت $f(y+5)-f(y)$ محاسبه می‌شود. سه آستانه مختلف برای تعیین دره‌ها و قله‌های منحنی در نظر گرفته می‌شود. نقطه‌ای که شیب آن بیشتر از اولین آستانه باشد بعنوان دره مربوط به پلک پایین شناخته می‌شود. با ادامه دادن آنالیز منحنی، نقاطی که شیب آنها کمتر از آستانه شماره دو باشند، مشخص می‌شوند. این نقاط در قله منحنی قرار دارند که نشان دهنده بخش میانی چشم می‌باشد. بطور مشابه با پلک پایین، پلک بالا در نقطه‌ای قرار دارد که شیب آن بیشتر از آستانه شماره سه باشد. در مقاله ارائه شده توسط Gan و همکارانش [۵]، مقدار عددی آستانه بیان نشده است.



(الف): تصویر باینری شده چهره

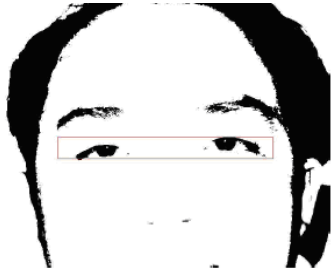


(ب): منحنی مربوط به تابع نگاشت جمع کننده در جهت افقی

شکل ۷-۵: تعیین مرزهای چشم

در روش پیشنهادی آنها مرزهای عمودی چشم با استفاده از تابع نگاشت جمع کننده در جهت عمودی تعیین می‌شود. موها و نقاط میانی چشمها از روی مقادیر بدست آمده برای این تابع مشخص می‌شوند و سپس یک نقطه بطو تصادفی در ناحیه بین این دو بخش بعنوان مرز عمودی چشم در نظر گرفته می‌شود. در روش Gan و همکارش [۵]، تشخیص چشمهای باز و بسته با استفاده مقدار تابع نگاشت جمع کننده در جهت افقی $IPF_h(y)$ انجام می‌شود. شکل ۸-۵ تصاویر مربوط به چشمهای باز و بسته را نشان می‌دهد.

همانطور که در شکل مشخص است در زمان باز بودن چشم، ناحیه چشم شامل مردمک و مژه‌هاست. در این حالت تعداد پیکسل‌های $IPF_h(y)$ که ارزش آنها صفر می‌باشد بسیار زیاد است و در نتیجه مقدار $IPF_h(y)$ نیز مقدار بزرگی خواهد بود. زمانی که چشم بسته است، ناحیه چشم فقط شامل مژه می‌شود و $IPF_h(y)$ کوچک خواهد بود. بنابراین با بررسی مقدار $IPF_h(y)$ در یک تصویر باز یا بسته بودن چشم مشخص می‌گردد.



(ب): تشخیص چشم‌های باز



(الف): تشخیص چشم‌های بسته

شکل ۵-۸: تشخیص باز یا بسته بودن چشم‌ها

در این روش، مطالعه تغییرات $IPF_h(y)$ در طول زمان برای تشخیص خواب آلودگی یا پلک زدن مورد استفاده قرار گرفت. اگر مقدار $IPF_h(y)$ به آرامی در طول زمان کاهش یابد مؤید خواب آلودگی است در صورتیکه تغییرات شدید $IPF_h(y)$ بیانگر پلک زدن می‌باشد.

۵-۲- الگوریتم پیشنهادی برای تعیین فاصله بین پلک‌ها

الگوریتم پیشنهاد شده در این تحقیق نیز با استفاده تابع نگاشت در جهت افقی به تعیین مکان پلک‌ها می‌پردازد. تعیین مکان پلک‌ها مهمترین قدم در تشخیص پلک زدن و تشخیص خواب آلودگی می‌باشد. کاملاً مشخص است که در تشخیص خواب آلودگی تصاویر چشم در شرایط گوناگونی از نظر نورپردازی تهیه می‌شوند و ممکن است سطح روشنایی تصویر تغییرات شدیدی داشته باشد. لذا مقاوم بودن الگوریتم نسبت به این مسئله از اهمیت زیادی برخوردار است. در روش پیشنهاد شده توسط Gan و همکارش [۵]، از تصویر باینری استفاده شد. ضعفی که در رابطه با این روش وجود دارد تعیین آستانه مناسب برای باینری نمودن تصویر اولیه می‌باشد. در صورتیکه مقدار کمی برای آستانه انتخاب شود، نویز زیادی در تصویر باینری بجا خواهد ماند و اگر مقدار تعیین شده برای آستانه مقدار بزرگی باشد ممکن است منجر به حذف اطلاعات مفید تصویر شود. بنابراین نتایج این روش کاملاً وابسته به آستانه مورد استفاده می‌باشند.

در الگوریتم پیشنهادی، برای اجتناب از این مشکل و غلبه بر تغییر نور محیط از گرادیان تصویر استفاده می‌شود.

همانطور که قبلاً نیز گفته شد سطح روشنایی پیکسل‌ها در ناحیه چشم بسیار پایین‌تر از نواحی همسایه‌اش می‌باشد و پلک‌ها مرز تغییر سطح روشنایی بین ناحیه چشم و پوست می‌باشند. بنابراین مقدار گرادیان در محل پلک‌ها، مقدار بزرگی خواهد بود. تصاویر ورودی در الگوریتم پیشنهادی برای تعیین فاصله بین پلک‌ها، تصاویر سطح خاکستری ناحیه چشم می‌باشند.

اولین مرحله از الگوریتم پیشنهادی محاسبه گرادیان تصویر ورودی می‌باشد. گرادیان تصویر f توسط رابطه ۷-۵ محاسبه می‌شود.

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} i_x + \frac{\partial f(x, y)}{\partial y} i_y \quad 7-5$$

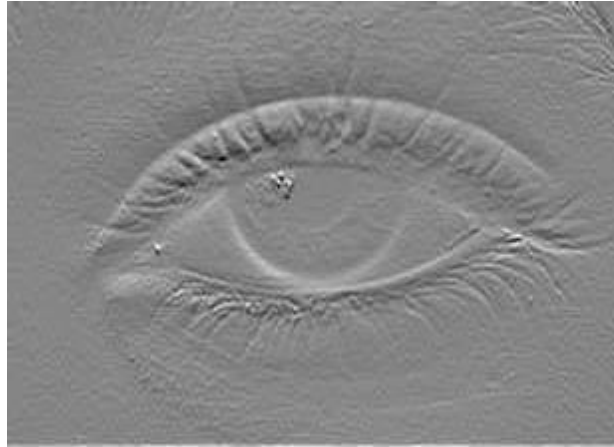
i_x و i_y در این رابطه، بردارهای یکه در جهت‌های x و y هستند. همانطور که رابطه ۷-۵ نشان می‌دهد، گرادیان یک کمیت برداری است. گرادیان تصویر در جهت محور افقی یعنی $\frac{\partial f(x, y)}{\partial x}$ ، تغییرات شدت روشنایی در راستای محور x را نشان می‌دهد و گرادیان تصویر در جهت محور عمودی $\frac{\partial f(x, y)}{\partial y}$ ، تغییرات شدت روشنایی در راستای محور y را نشان می‌دهد. دامنه گرادیان از رابطه ۸-۵ محاسبه می‌گردد و بیانگر ماکزیمم تغییرات شدت روشنایی در نقطه (x, y) می‌باشد.

$$|\nabla f(x, y)| = \sqrt{\left(\frac{\partial f(x, y)}{\partial x}\right)^2 + \left(\frac{\partial f(x, y)}{\partial y}\right)^2} \quad 8-5$$

بدلیل اینکه هدف ما در الگوریتم پیشنهادی، تعیین مکان پلک‌ها بوده‌است، بنابراین کفایت تغییر شدت روشنایی در راستای محور عمودی بررسی شود. بنابراین گرادیان تصویر در جهت عمودی محاسبه می‌شود. شکل ۹-۵ تصویر یک چشم را به همراه تصویر گرادیان در جهت عمودی و تصویر اندازه گرادیان در جهت عمودی نشان می‌دهد. سطح روشنایی پیکسل‌ها در تصویر شکل ۹-۵ (ج)، نشان دهنده دامنه گرادیان در جهت عمودی برای آن پیکسل می‌باشد. با دقت در تصویر در می‌یابیم سطح روشنایی پیکسل‌ها در مکان پلک بالا و پلک پایین بیشترین مقدار را دارد. بنابراین با جستجوی این نقاط در تصویر مربوط به اندازه گرادیان مکان پلک‌ها تعیین می‌شود.



(الف): تصویر سطح خاکستری چشم



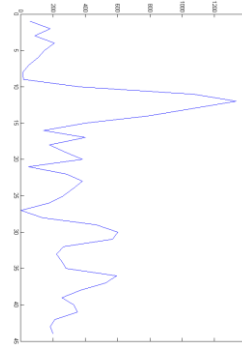
(ب): تصویر گرادیان در جهت عمودی



(ج): تصویر دامنه گرادیان در جهت عمودی

شکل ۵-۹: شناسایی مرزهای افقی چشم

با توجه به مطالب گفته شده، قدم دوم در الگوریتم پیشنهادی جستجو به منظور یافتن نقاطی است که اندازه گرادیان در جهت عمودی برای آنها مقدار بزرگی دارد. اما یافتن مختصات نقاط به تنهایی و بدون در نظر گرفتن سطح روشنایی پیکسل‌های همسایه‌اش برای تعیین مکان پلک‌ها کافی نیست. شکل ۵-۱۰(الف) تصویر دیگری از اندازه گرادیان در جهت عمودی را نشان می‌دهد. همانطور که در شکل دیده می‌شود، سطرهایی از تصویر که مربوط به مکان پلک‌ها می‌باشند، بیشترین فراوانی را از نظر تعداد پیکسل‌های با شدت روشنایی بالا دارند. تابع نگاشت جمع کننده در جهت افقی، قادر است فراوانی این پیکسل‌ها را در هر سطر محاسبه کند. شکل ۵-۱۰(ب) مقدار $IPF_h(\gamma)$ را برای شکل نشان می‌دهد.

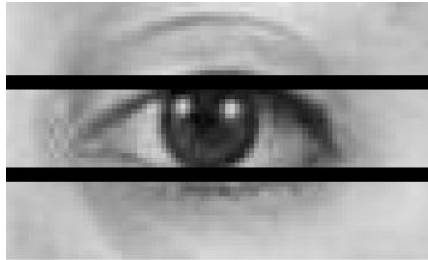


(الف): اندازه گرادیان تصویر در جهت عمودی

(ب): تابع نگاشت جمع کننده در جهت افقی

شکل ۵-۱۰: پاسخ تابع نگاشت عمودی در مکان پلکها

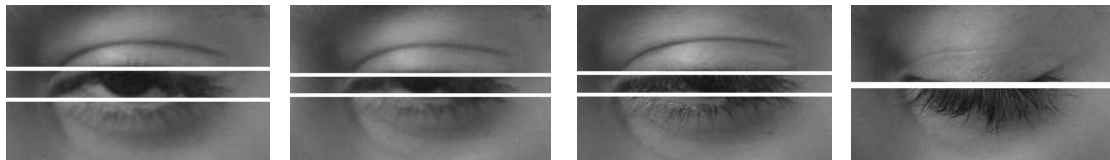
در آخرین مرحله از الگوریتم پیشنهادی، منحنی $IPF_h(y)$ به دو قسمت مساوی تقسیم می‌شود. سطری از تصویر که در نیمه بالای این منحنی بیشترین مقدار را داشته باشد بعنوان پلک بالا در نظر گرفته می‌شود. و سطری که در نیمه پایین این منحنی بیشترین مقدار را داشته باشد مکان پلک پایین را مشخص می‌کند. شکل ۵-۱۱ تصویر سطح خاکستری چشم را نشان می‌دهد که مکان پلکها در آن توسط الگوریتم پیشنهادی مشخص شده‌اند.



شکل ۵-۱۱: تشخیص مکان پلکها

اختلاف مکان دو پلک، فاصله آنها بر حسب پیکسل می‌باشد. افت این فاصله در طول زمان نشان‌دهنده خواب آلودگی می‌باشد و تغییر ناگهانی این فاصله در پلک زدن اتفاق می‌افتد.

تصاویر مربوط به یک دنباله ویدئویی همراه با نتایج بدست آمده در هر مرحله از الگوریتم پیشنهادی در شکل ۵-۱۲ تا نشان داده شده است.

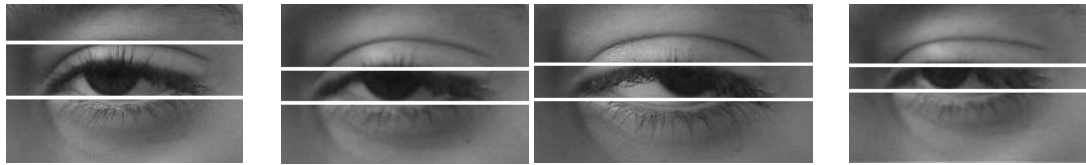


(۴)

(۳)

(۲)

(۱)

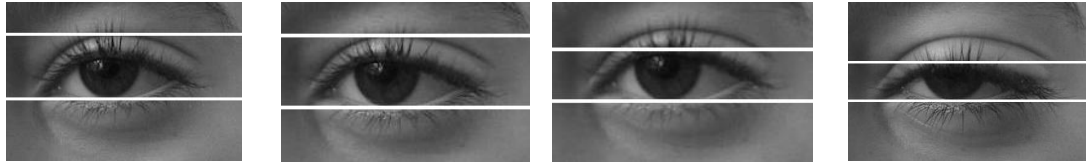


(۸)

(۷)

(۶)

(۵)

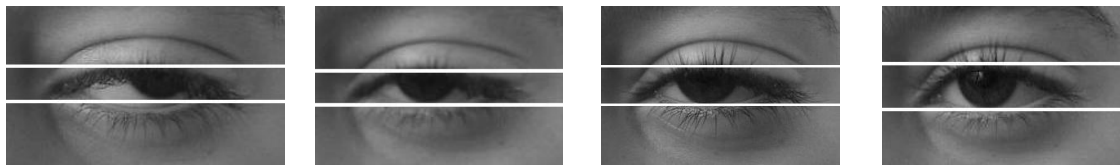


(۱۲)

(۱۱)

(۱۰)

(۹)

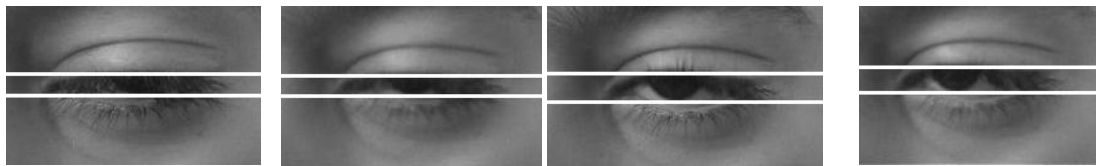


(۱۶)

(۱۵)

(۱۴)

(۱۳)



(۲۰)

(۱۹)

(۱۸)

(۱۷)



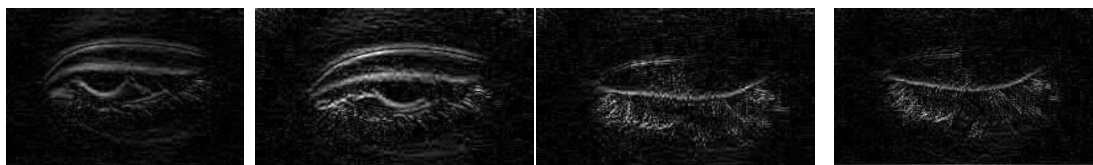
(۲۴)

(۲۳)

(۲۲)

(۲۱)

شکل ۵-۱۲: نتایج الگوریتم ردیابی چشم در فرایند پلک زدن

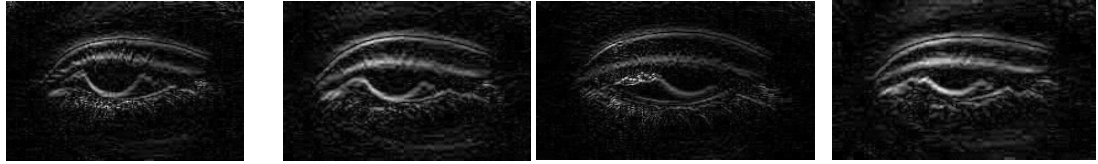


(۴)

(۳)

(۲)

(۱)

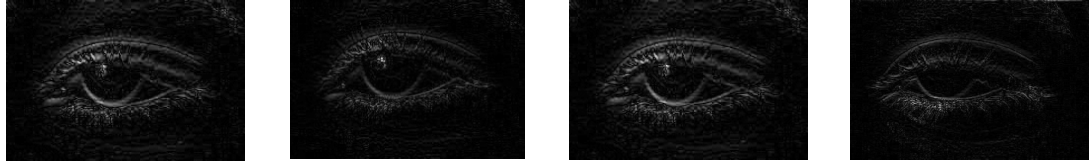


(۸)

(۷)

(۶)

(۵)

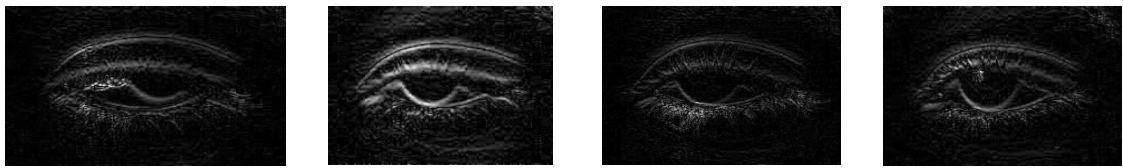


(۱۲)

(۱۱)

(۱۰)

(۹)

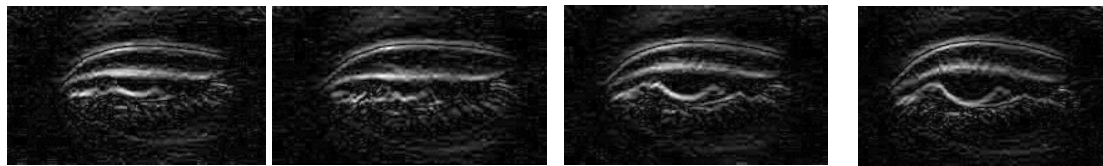


(۱۶)

(۱۵)

(۱۴)

(۱۳)

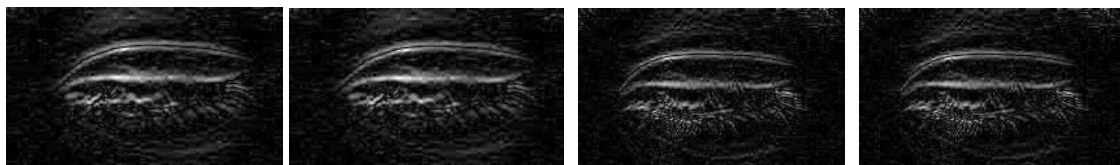


(۲۰)

(۱۹)

(۱۸)

(۱۷)



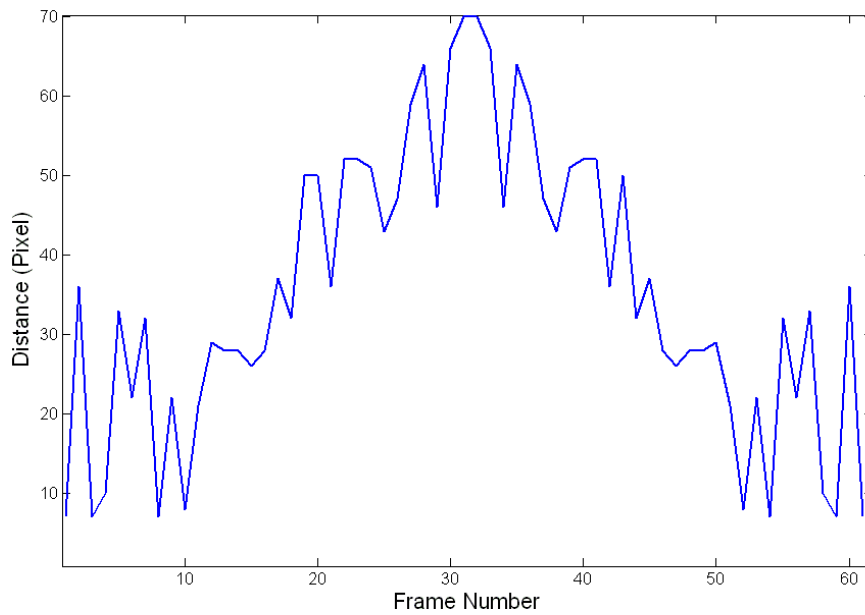
(۲۴)

(۲۳)

(۲۲)

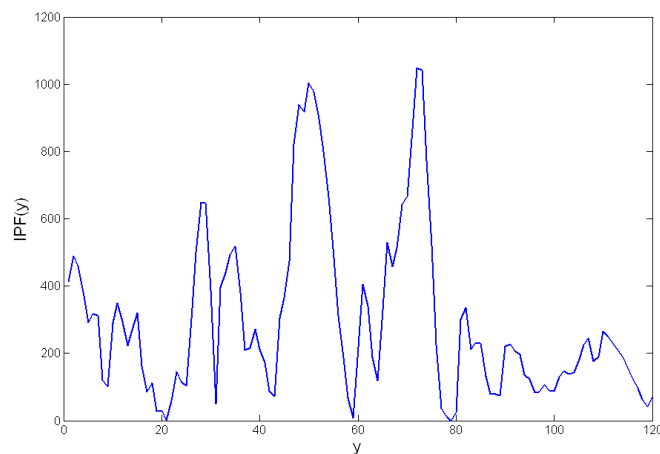
(۲۱)

شکل ۵-۱۳: تصاویر اندازه گرا دیان روشنایی در جهت عمودی

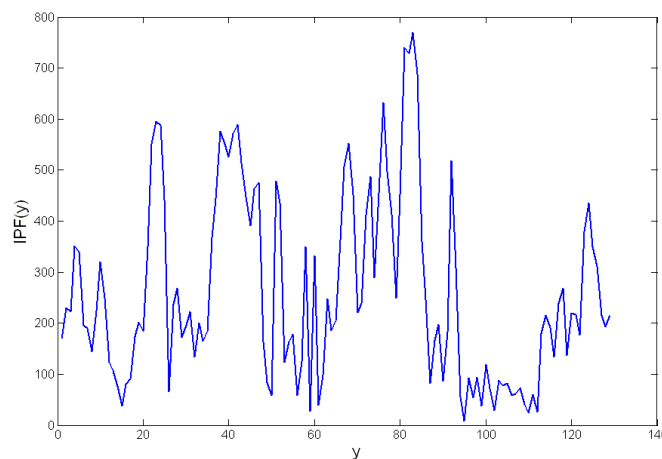


شکل ۵-۱۴: منحنی نمایش دهانه فاصله پلک‌ها در پلک زدن

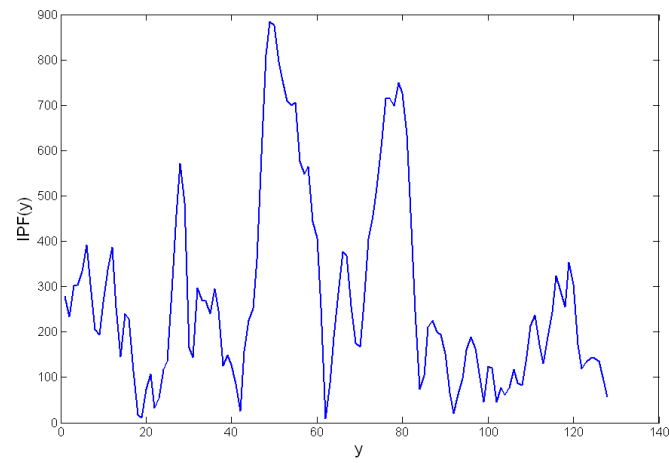
در شکل‌های ۵-۱۵ تا ۵-۱۷ تابع نگاشت جمع کننده $IPF(y)$ در جهت افقی برای سه فریم که در شکل ۵-۱۲ با شماره‌های ۴، ۱۰ و ۱۳ مشخص شده‌اند، دیده می‌شود.



شکل ۵-۱۵: منحنی تابع نگاشت جمع کننده در جهت افقی برای شکل شماره ۴



شکل ۵-۱۶: منحنی تابع نگاشت جمع کننده در جهت افقی برای شکل شماره ۱۰



شکل ۵-۱۷: منحنی تابع نگاشت جمع کننده در جهت افقی برای شکل شماره ۱۳

فصل ۶- نتایج تجربی

این بخش نتایج بدست آمده توسط الگوریتم‌های پیشنهادی در این تحقیق را ارائه می‌کند. در اولین قدم پایگاه‌های داده مورد استفاده در این تحقیق معرفی شده سپس برای ارزیابی عملکرد سه الگوریتم تشخیص چشم، ردیابی چشم و تعیین فاصله بین پلک‌ها نتایج حاصله توسط آنها بطور جداگانه بررسی می‌گردد.

۶-۱- پایگاه داده

با توجه به اینکه پایگاه داده و شرایط اعمال شده در فرایند تهیه آن، نقش بسزایی در نتایج الگوریتم‌های بینایی ماشین و پردازش تصویر دارد، انتخاب پایگاه داده مناسب نیازمند دقت و توجه زیادی است. پایگاه‌های داده مورد استفاده در این تحقیق را می‌توان به دو بخش تقسیم نمود:

۱. بخش مورد نیاز برای آموزش سیستم

۲. بخش مربوط به آزمون الگوریتم‌های پیشنهادی

۶-۱-۱- داده‌های آموزشی

بطور کلی داده‌ها یا نمونه‌های مورد نیاز برای آموزش یک کلاسه‌بند شامل دو گروه یعنی نمونه‌های مثبت و نمونه‌های منفی می‌شود. در این تحقیق نمونه‌های مثبت تصاویر چشم می‌باشند که به دو دسته تصاویر چشم راست و چشم چپ تقسیم شده‌اند و نمونه‌های منفی نیز تصاویر غیر چشم هستند. این نمونه‌ها از قسمت‌های دیگر چهره، که شامل چشم‌ها نمی‌شود جدا شده‌اند. همانطور که در فصل ۳ نیز بیان شد، جمع‌آوری داده‌های آموزشی به صورت بالا منجر به ایجاد سه گروه چشم راست، چشم چپ و غیر چشم می‌گردد که از آنها برای آموزش الگوریتم و تخمین مدل آماری سه گروه فوق استفاده می‌شود.

ما بدنبال مدل آماری هستیم تا پراکندگی ویژگی‌های توپولوژیکی را برای سه کلاس چشم راست، چشم چپ و غیر چشم بیان نماید. آشکارسازی چشم‌ها در تصویر مستقیماً از این مدل تأثیر می‌پذیرد. از طرفی ایجاد یک مدل مناسب در گروه انتخاب درست نمونه‌های آموزشی می‌باشد.

در ادامه به بررسی نکاتی می‌پردازیم که در تهیه پایگاه داده آموزشی رعایت شده‌اند.

همانطور که گفته شد تصاویر موجود در سه گروه مذکور در پایگاه داده آموزشی از بخش‌های مختلف تصویر چهره جدا می‌شوند. جدا کردن این نواحی (چشم راست، چشم چپ و غیر چشم) به شیوه ذیل انجام می‌شود. با توجه به اینکه نقاط کاندید برای چشم دارای برجسب توپوگرافیکی چاله که مکان مردمک را نشان می‌دهد می‌باشند، لذا جهت آموزش سیستم تصاویر چشم‌ها را بنحوی برش می‌دهیم که چاله در مرکز آنها قرار گیرد. در خصوص آموزش نواحی غیر چشم، نیز بخش‌هایی از تصویر چهره انتخاب می‌شوند که دارای برجسب چاله در مرکز باشند.

در اولین مرحله ساختار توپوگرافیک تصویر چهره استخراج می‌شود و چاله‌ها در تصویر برچسب خورده مشخص می‌شوند. مشابه الگوریتم تست، پس از دسته بندی و حذف کاندیدهای نادرست، لیستی از چاله‌ها بصورت دوبه‌دو تهیه می‌شود و یک ناحیه مستطیل شکل با ابعاد $0.6d \times 0.3d$ در جهت خط واصل بین دو چاله جدا می‌شود (d طول خط واصل دو چاله (مردمک) است). در پیاده‌سازی روش Wang و همکارش [۴۷] می‌بایست تمام نواحی جدا شده از تصویر که مستقیماً بعنوان بردار ویژگی هر ناحیه استفاده می‌شود، ابعاد یکسانی داشته باشند. لذا در تهیه داده‌های آموزشی مورد نیاز این روش، ابعاد ناحیه ثابت و برابر با 80×40 پیکسل در نظر گرفته شد. این ابعاد بصورتی انتخاب شدند که ناحیه چشم را بطور کامل در تمامی تصاویر چهره در برگیرد. پس از جدا کردن نمونه‌های آموزشی، آنها را در سه گروه چشم چپ و راست و غیر چشم دسته بندی می‌کنیم. دقت در تصاویر چهره نشان می‌دهد که اکثر چاله‌ها در تصویر توپوگرافیک علاوه بر مراکز مردمک در قسمت‌های خاصی نظیر سوراخ‌های بینی، گوشه‌های لب، موها و ابروها ظاهر می‌شوند.

به منظور مقاوم سازی الگوریتم تشخیص چشم‌ها در برابر تغییراتی نظیر سن، نژاد و جنسیت افراد سعی شده‌است تا حتی‌الامکان تصاویر چهره مورد استفاده برای آموزش سیستم محدوده وسیعی از تغییرات عوامل بالا را پوشش دهد. لازم به ذکر است که بدلیل اینکه ساختار توپوگرافیک تصویر تا حد زیادی نسبت به تغییرات شرایط نورپردازی و تغییر کنتراست مقاوم است، نیازی به تهیه تصاویر آموزشی در شرایط مختلف نورپردازی و کنتراست نمی‌باشد.

در این تحقیق دو مجموعه آموزشی تهیه شد که مجموعه اول برای پیاده‌سازی الگوریتم Wang و همکارش [۴۷] مورد استفاده قرار گرفت و مجموعه دوم نیز به منظور آموزش الگوریتم پیشنهادی در این تحقیق تهیه شد.

در مجموعه اول، ۱۰۰ تصویر برای گروه چشم راست، ۱۰۰ تصویر برای گروه چشم چپ و ۱۰۰۰ تصویر نیز برای گروه غیر چشم تهیه شد. در مجموعه دوم تعداد تصاویر گروه چشم راست ۴۰۱ عدد، تعداد تصاویر گروه چشم چپ ۴۵۲ عدد و تعداد تصاویر گروه غیر چشم نیز ۱۳۶۵ انتخاب شد. نکته قابل ذکر در رابطه با تصاویر مجموعه آموزشی دوم امکان استفاده از تصاویر با ابعاد متفاوت می‌باشد. این بدلیل اینکه در روش پیشنهادی برخلاف روش Wang و همکارش [۴۷] ابعاد بردار ویژگی مستقل از ابعاد تصویر می‌باشد، میسر می‌گردد. بخشی از پایگاه داده XM2VTS شامل تصاویر ۱۳۹ فرد در تهیه پایگاه‌های آموزشی مورد استفاده قرار گرفت.

۶-۱-۲- پایگاه داده مربوط به تست الگوریتم

به منظور ارزیابی هرچه بهتر روشهای پیشنهادی و مقایسه آن با روش Wang و همکارش [۴۷]، نیاز به پایگاه داده‌ای شامل تصاویر چهره متنوع داری می‌تصاویر موجود در این پایگاه داده می‌بایست در شرایط متنوعی از لحاظ تصویری برداری، چرخش سر، حالات چهره، وجود عینک و ... تهیه شده باشند. پایگاه‌های داده مورد استفاده در این بخش به دو گروه تقسیم می‌شود:

۱. پایگاه داده استفاده شده برای تست الگوریتم تشخیص چشم

۲. پایگاه داده استفاده شده برای تست الگوریتم ردیابی چشم و الگوریتم تعیین فاصله بین پلکها

مجموعه‌ای که برای تست الگوریتم تشخیص چشم مورد استفاده قرار می‌گیرد بخشی از پایگاه داده XM2VTS که در تهیه تصاویر آموزشی نقش ندارد، می‌باشد. پایگاه داده XM2VTS شامل تصاویر ۳۷۱ فرد می‌باشد که برای هر شخص در این پایگاه ۸ تصویر در حالات گوناگون و زوایای مختلف چرخش سر وجود دارد. کلیه تصاویر موجود برای ۲۳۱ فرد در این پایگاه داده به منظور تست الگوریتم مورد استفاده قرار گرفت. تصاویر موجود در این پایگاه داده تصاویر رنگی می‌باشند که پس از تبدیل به تصاویر سطح خاکستری مورد استفاده قرار گرفتند.

برای آزمایش الگوریتم ردیابی چشم‌ها مجموعه‌ای از ۱۰۰ دنباله ویدئویی از افراد مختلف تهیه شد. هریک از دنباله‌های ویدئویی تهیه شده شامل ۴۵۰ تا ۵۰۰ فریم می‌باشد و سعی شده‌است تا حد امکان فاصله سر با دوربین، چرخش سر در زوایای گوناگون و مکان چشم‌ها در محدوده وسیعی تغییر کند. همچنین تغییر در نورپردازی و استفاده از عینک نیز در تهیه این تصاویر در نظر گرفته شده‌است. از همین پایگاه داده برای تست الگوریتم پیشنهادی برای تعیین فاصله بین پلکها نیز استفاده شد.

۶-۲- نتایج

در این بخش به بررسی نتایج حاصل از اعمال الگوریتم‌های پیشنهادی بر روی پایگاه‌های داده جمع‌آوری شده می‌پردازیم

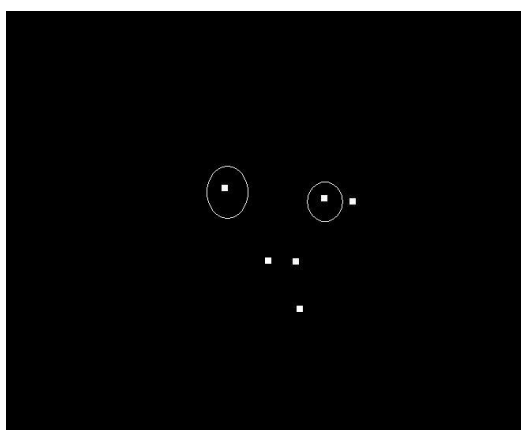
۶-۲-۱- نتایج الگوریتم تشخیص چشم

در این تحقیق ابعاد پنجره فیلتر هموار ساز گوسی 17×17 و پارامتر انحراف معیار آن $\delta = 4.7$ تعیین شد. ابعاد پنجره فیلتر مشتق‌گیر مبتنی بر چندجمله‌ای چبی‌چف نیز 5×5 انتخاب شد. این مقادیر با آزمایش‌های متعدد و مقایسه نتایج بدست آمد.

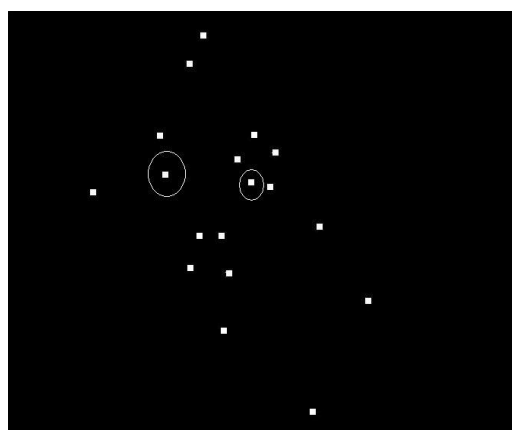
نتایج بدست آمده نشان می‌دهد انتخاب مقدار بالا برای ابعاد پنجره و انحراف معیار در فیلتر هموارساز گوسی موجب حذف بسیاری از چاله‌ها در تصویر توپوگرافیک می‌شود. از طرفی انتخاب مقدار کم برای این پارامترها در فیلتر منجر به تولید چاله‌های زیادی در تصویر می‌شود که بسیاری از آنها متأثر از نویز و تغییرات روشنایی اندک محیطی می‌باشند. اما آزمایشات گوناگون نشان می‌دهد چاله مربوط به مردمک در تصاویر توپوگرافیکی از پایداری بسیار بالایی نسبت به تغییر پارامترهای این فیلتر برخوردار است. شکل ۶-۱ مثالی از نتایج بدست آمده در این رابطه را نشان می‌دهد. تصاویر باینری، توزیع برجسب‌های چاله در تصویر توپوگرافیکی را نشان می‌دهند. همانطور که در شکل ۶-۱ دیده می‌شود با تغییر پارامترهای فیلتر هموارساز گوسی در محدوده وسیع، چاله‌های مربوط به مردمک همچنان در تصویر توپوگرافیکی حضور دارند. در آخرین تصویر، فقط چاله‌های مربوط به مردمک باقی مانده‌اند.



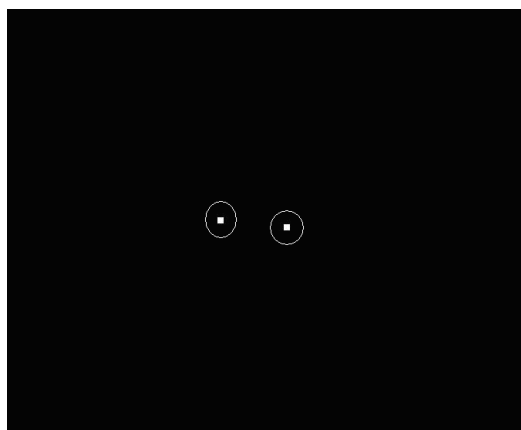
(الف) : تصویر آزمون



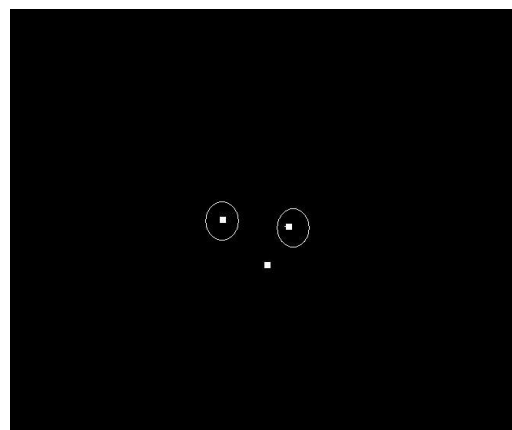
$\delta = 4.7, N = 17$



$\delta = 3, N = 17$



$\delta = 15, N = 25$



$\delta = 10, N = 25$

(ب) : آشکارسازی چاله‌ها در تصویر توپوگرافیکی

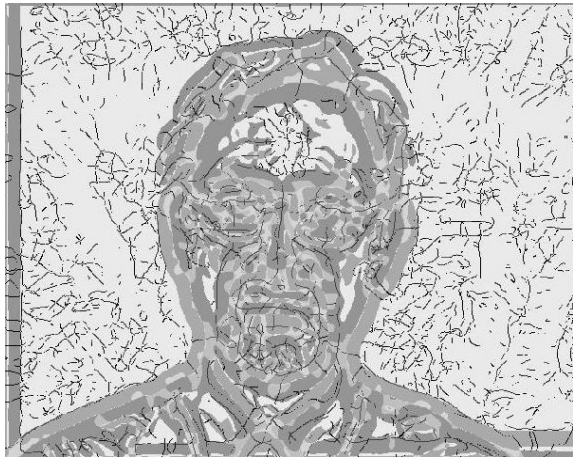
شکل ۶-۱: پایداری چاله‌های مربوط به مردمک در اثر تغییر پارامترهای فیلتر هموارساز گوسی

با توجه به اینکه در الگوریتم تشخیص چشم ابتدا نقاط کاندیدای مردمک در تصاویر آزمون تعیین می‌شوند، لذا توانایی الگوریتم در تعیین چاله‌های مربوط به مردمک از اهمیت زیادی برخوردار است. بهمین دلیل نتایج الگوریتم در مرحله تشخیص نقاط کاندید مورد بررسی و تحلیل قرار گرفت. برای این منظور از

۱۴۰۲ تصویر چهره استفاده گردید. اولین مرحله در تعیین نقاط کاندید مردمک پس از تعیین ساختار توپوگرافیک تصویر آزمون صورت می‌گیرد. نتایج این مرحله حاکی از آن بود که الگوریتم پیشنهادی در کلیه تصاویر آزمون، چاله‌های مربوط به مردمک را بطور صحیح آشکار می‌سازد. توانایی الگوریتم در مرحله دوم که مربوط به گروه‌بندی و حذف کاندیدهای نادرست می‌باشد نیز مورد بررسی قرار گرفت. نتایج این مرحله در مورد ۶۸ تصویر از تصاویر آزمون با خطا مواجه شد بنحویکه یکی از دو چاله یا هر دو چاله مربوط به مردمک چشم‌ها بعنوان کاندید نادرست حذف شدند.

لازم به ذکر است که تغییر پارامترهای مربوط به فیلتر هموارساز گوسی و ابعاد پنجره فیلتر مشتق‌گیر ساختار توپوگرافیک تصویر را تحت تأثیر قرار می‌دهند. آزمایشات گوناگون نشان می‌دهد که مقدار این پارامترها باید با توجه به ابعاد تصویر ورودی و فاصله سر با دوربین انتخاب شود.

شکل‌های ۲-۶ و ۳-۶ تغییر نتیجه در اثر تغییر ابعاد پنجره فیلتر مشتق‌گیر (n) و تغییر پارامترهای فیلتر هموارساز گوسی را نشان می‌دهند:



تصویر برجسب خورده توپوگرافیکی



تصویر آزمون

شکل ۲-۶ (الف): نتیجه بدست آمده برای $n=5$



تصویر برجسب خورده توپوگرافیکی



تصویر آزمون

(ب) : نتیجه بدست آمده برای $n=9$

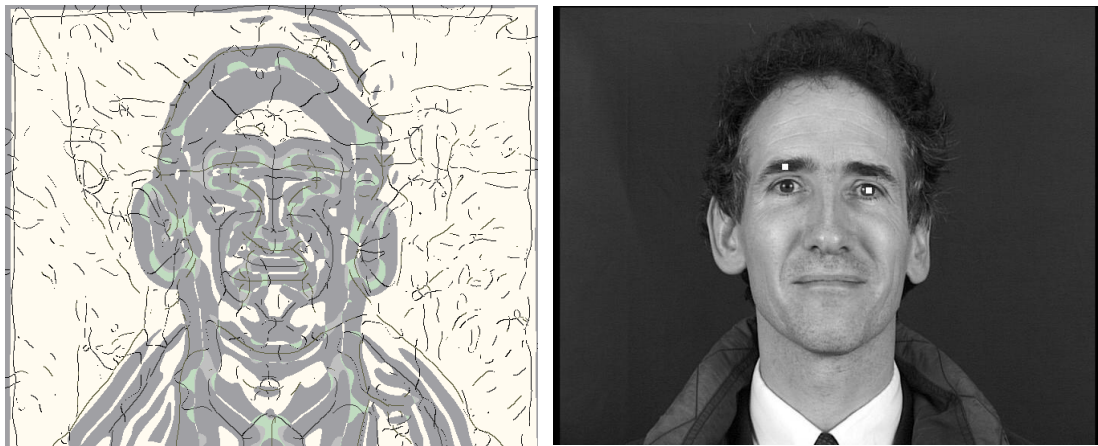
شکل ۶-۲: مقایسه نتایج در اثر تغییر ابعاد پنجره فیلتر مشتق گیر



تصویر برجسب خورده توپوگرافیکی

تصویر آزمون

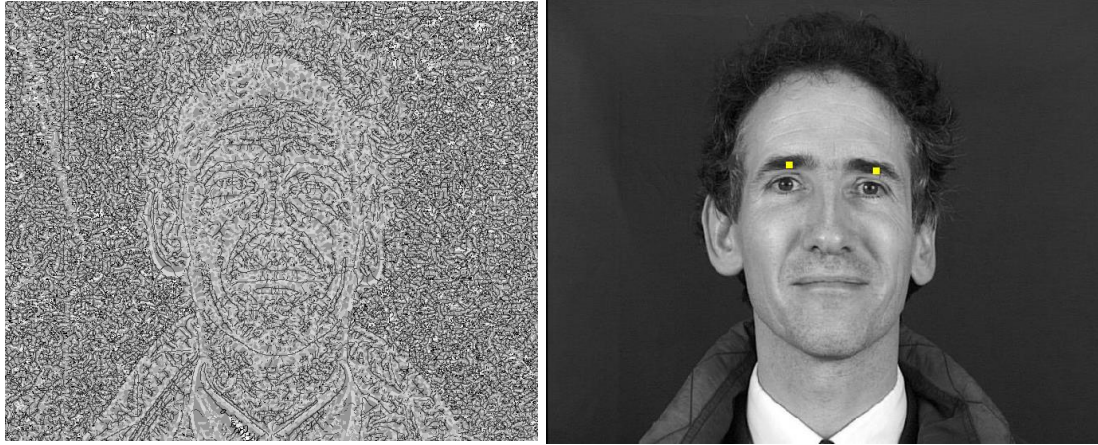
شکل ۶-۳(الف) : نتیجه بدست آمده برای $\delta = 4.7, N = 17$



تصویر برجسب خورده توپوگرافیکی

تصویر آزمون

(ب) : نتیجه بدست آمده برای $\delta = 6.5, N = 170$



تصویر برجسب خورده توپوگرافیکی

تصویر آزمون

(ج) : نتیجه بدست آمده برای $\delta = 4.7, N = 5$

شکل ۳-۶ : مقایسه نتایج در اثر تغییر پارامترهای فیلتر هموارساز گوسی

به منظور مقایسه کارایی الگوریتم پیشنهادی با روش ارائه شده توسط Wang و همکارش [۴۷]، عملکرد دو الگوریتم بر روی ۱۰۰ تصویر از پایگاه داده مورد بررسی قرار گرفت. برای این منظور هر دو الگوریتم با استفاده نمونه‌های آموزشی تهیه شده در مجموعه ۱، آموزش دیدند.

جدول زیر نتایج بدست آمده توسط دو الگوریتم را نشان می‌دهد:

جدول ۱-۶ : مقایسه نرخ آشکارسازی برای روش Wang و همکارش [۴۷] و الگوریتم پیشنهادی

نرخ آشکارسازی	تشخیص خطا در مورد هر دو چشم	تشخیص مکان چشم چپ	تشخیص مکان چشم راست	تشخیص مکان دو چشم	
٪۸۲	۶	۵	۷	۸۲	نتایج مربوط به الگوریتم Wang, Yin
٪۹۰	۳	۳	۴	۹۰	نتایج مربوط به الگوریتم پیشنهادی

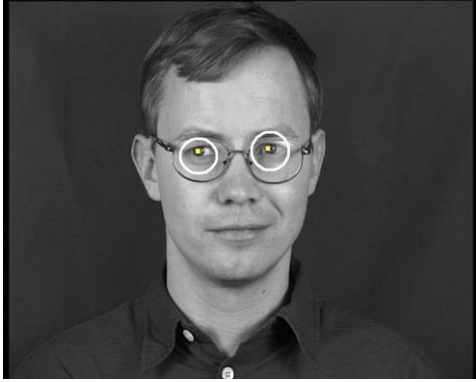
مقایسه دو روش نشان می‌دهد که روش پیشنهادی در مقایسه با روش ارائه شده توسط Wang و همکارش [۴۷] از سرعت بالاتری برخوردار می‌باشد. همانطور که گفته شد تصاویر آموزشی مربوط به این الگوریتم دارای ابعاد 80×40 پیکسل می‌باشد که منجر به تولید بردار ویژگی و بردار میانگین با 3200×3200 عنصر و ماتریس کواریانس با ابعاد 3200×3200 می‌شود. بدیهی است که ابعاد بسیار بالا متناظر با صرف زمان زیادی برای تعیین پاسخ می‌باشد. در صورتیکه در روش پیشنهادی بردار ویژگی بکار رفته تنها ۳۵

عنصر دارد و محاسبه ممان‌ها و توابع متشکل از این ممان‌ها که عناصر بردار ویژگی هستند نیز بار محاسباتی بسیار کمی دارد. جدول ۲-۶ مقایسه سرعت دو الگوریتم را نشان می‌دهد:
 جدول ۲-۶: مقایسه سرعت برای روش Wang و همکارش [۴۷] و الگوریتم پیشنهادی

زمان اجرا	سیستم	روش
۱ دقیقه و ۴۱ ثانیه	MATLAB 6.0, Core(TM)2 ,1.83GHZ,1.GB RAM	الگوریتم Wang, Yin
۲۰ ثانیه	MATLAB 6.0, Core(TM)2 ,1.83GHZ,1.GB RAM	الگوریتم پیشنهادی

شکل ۴-۶ تعدادی از نتایج بدست آمده از پیاده‌سازی روش Wang و همکارش را نشان می‌دهد. این نتایج حاصل آموزش الگوریتم‌ها توسط نمونه‌های آموزشی حاضر در مجموعه اول می‌باشد.

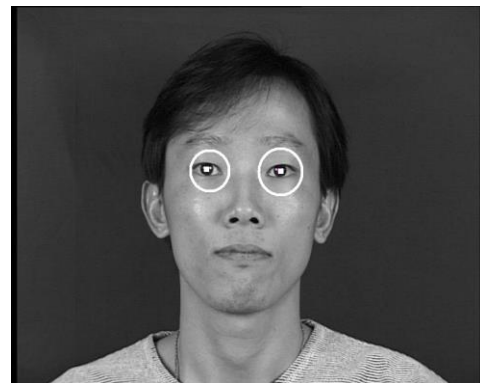
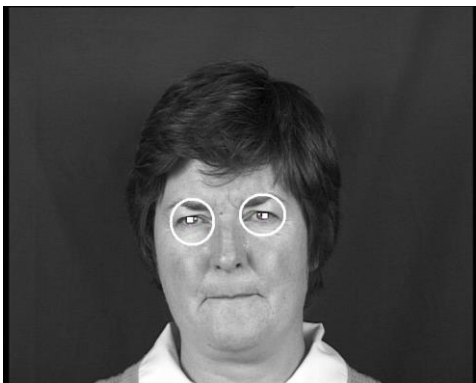




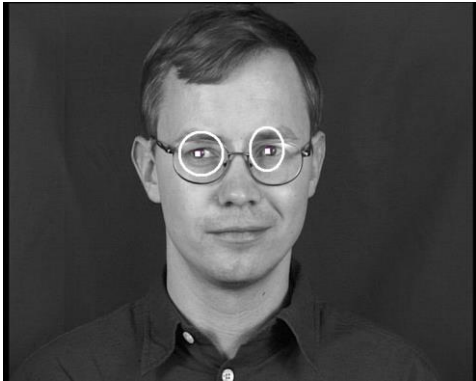
شکل ۴-۶: نتایج حاصل از پیاده‌سازی الگوریتم Wang و همکارش [۴۷]

در روش Wang و همکارش [۴۷]، شرط حضور برجسب ناودانی بیرون^۱ بین دو نقطه کاندید مردمک مورد استفاده قرار نگرفت. به همین دلیل الگوریتم ارائه شده توسط آنها در موارد زیادی سوراخ‌های بینی را چشم تشخیص داده است (شکل ۴-۶).

شکل ۵-۶ نتایج الگوریتم پیشنهادی برای تصاویر شکل ۴-۶ را نشان می‌دهد:



¹ Ridge



شکل ۵-۶: نتایج الگوریتم پیشنهادی

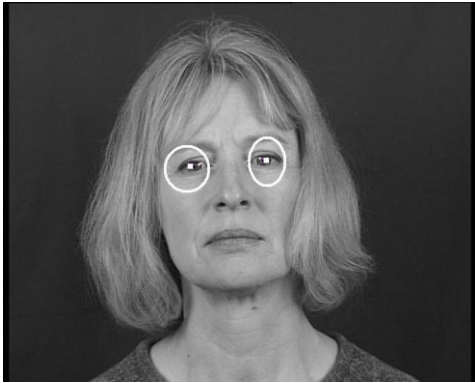
در ادامه به بررسی نتایج نهایی الگوریتم پیشنهادی می‌پردازیم. همانطور که گفته شد تصاویر با ابعاد متفاوت نیز در آموزش الگوریتم پیشنهادی قابل استفاده می‌باشند. جدول زیر نتایج بدست آمده توسط الگوریتم پیشنهادی را نشان می‌دهد:

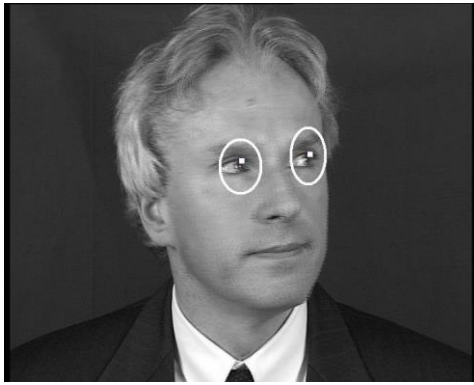
جدول ۳-۶: نتایج الگوریتم پیشنهادی

تعداد تصاویر آزمون	تشخیص مکان هر دو چشم	تشخیص مکان چشم راست	تشخیص مکان چشم چپ	تشخیص خطا در مورد هر دو چشم	نرخ آشکارسازی
۱۳۳۴	۱۲۳۷	۲۲	۳۷	۳۷	٪۹۲.۸

تعدادی از نتایج بدست آمده توسط الگوریتم پیشنهادی در شکل ۶-۶ نشان داده شده‌است.





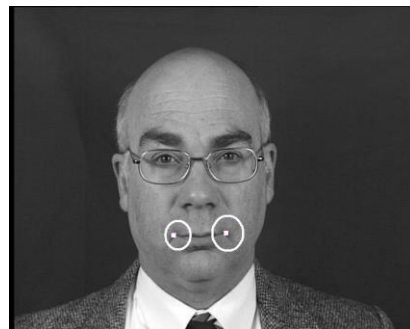


شکل ۶-۶: تعدادی از نتایج بدست آمده توسط الگوریتم پیشنهادی

توجه به تصاویری که نتیجه تشخیص چشم در آنها با خطا مواجه شده است، نشان می‌دهد بیشترین مناطقی از چهره که اشتباهاً چشم انتخاب می‌شوند، ابروها، گوشه‌های لب و قسمتی از فریم عینک هستند. مراجعه به تصویر چهره در حوزه توپوگرافیک بیانگر این واقعیت می‌باشد که ساختار توپوگرافیک این نواحی شبیه به چشم می‌باشد. با استفاده از برچسب سرایشی بدون انحناء و ناودانی بیرون در مرحله حذف و دسته‌بندی کاندیدها، امکان خطای الگوریتم در مناطق مربوط به سوراخ‌های بینی و موها تقریباً به صفر رسیده است. علاوه بر موارد مذکور چرخش عمقی سر در زوایایی که منجر به تغییر مدل چشم در حوزه توپوگرافیک می‌شود نتایج نادرستی تولید می‌کند. تعدادی از تصاویر آزمون که نتیجه الگوریتم پیشنهادی در رابطه با آنها با خطا مواجه شده است در شکل ۶-۷ دیده می‌شوند:



شکل ۶-۷: تعدادی از نتایج خطا در الگوریتم پیشنهادی





شکل ۶-۷: تعدادی از نتایج خطا در الگوریتم پیشنهادی

۶-۲-۲- نتایج الگوریتم ردیابی چشم

نتایج حاصل از اعمال الگوریتم ردیابی چشم بر روی پایگاه داده تهیه شده از دنباله‌های ویدئویی نشان داد که عوامل زیر در دنبال کردن مکان چشم‌ها در یک رشته ویدئویی باعث خطا می‌شوند:

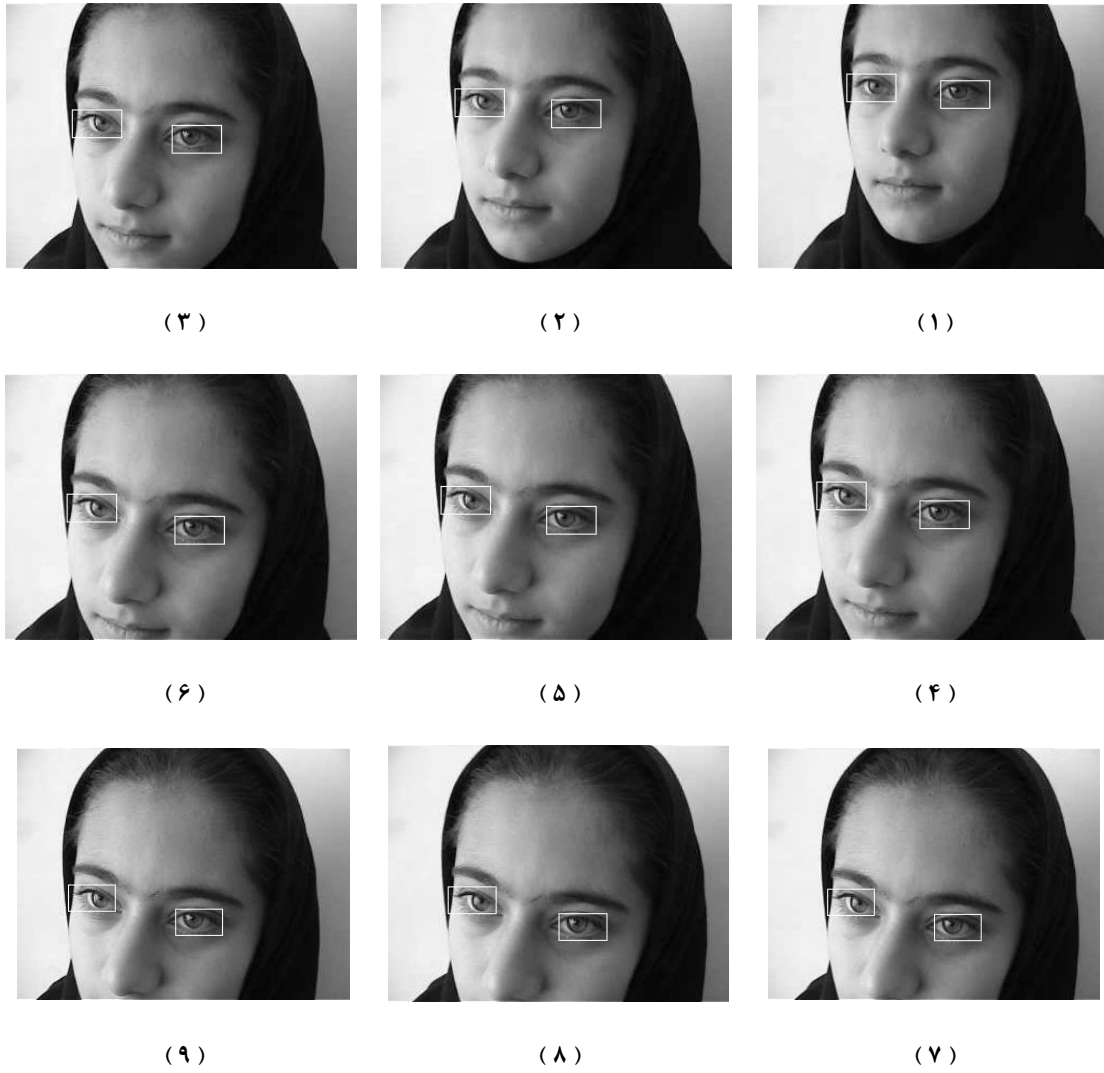
۱. چرخش عمقی سر به اندازه‌ای که باعث ناپدید شدن چشم‌ها گردد.

۲. بسته شدن چشم‌ها که در پلک زدن اتفاق می‌افتد.

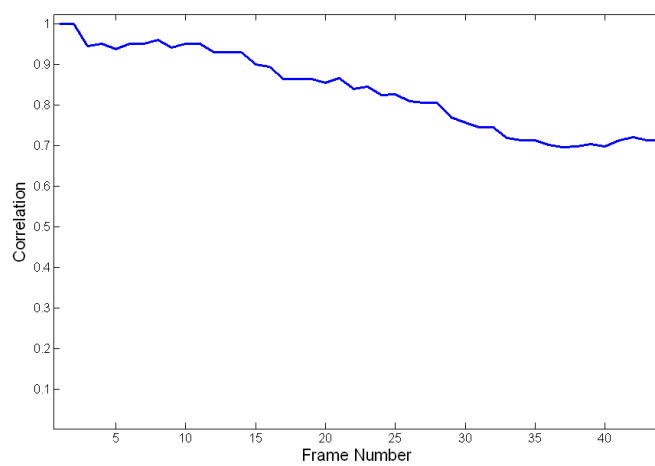
۳. فاصله سر تا دوربین بسیار زیاد یا بسیار کم شود.

نتایج بدست آمده توسط الگوریتم ردیابی چشم‌ها در این بخش ارائه شده‌است. به منظور ارزیابی موفقیت سیستم دنبال کننده چشم از همبستگی متقابل بین الگوی اولیه برای هر کدام از چشم‌ها و پنجره بدست آمده برای آن چشم در فریم مورد بررسی استفاده می‌گردد. نتایج ارائه شده در این بخش همراه با منحنی تغییرات تابع همبستگی متقابل در طول زمان می‌باشند. این منحنی برای چشم راست و چشم چپ بطور جداگانه تهیه شده‌است. محور افقی در منحنی تغییرات تابع همبستگی متقابل، شماره فریم‌های مورد بررسی را نشان می‌دهد و محور عمودی نیز مقدار تابع همبستگی بین الگوی اولیه و مکان چشم را در فریم مربوطه مشخص می‌کند. همانطور که در فصل ۴ گفته شد، افت مقدار تابع همبستگی متقابل با گذشت زمان، بیانگر افزایش خطا در نتیجه ردیابی چشم‌ها می‌باشد. با کاهش مقدار تابع به کمتر از آستانه تعیین شده برای همبستگی متقابل، خطا در الگوریتم ردیابی غیر قابل چشم‌پوشی خواهد بود. در این مرحله نیاز به تشخیص مجدد مکان چشم‌ها توسط الگوریتم تشخیص چشم می‌باشد. در الگوریتم پیشنهادی مقدار آستانه برای تابع همبستگی، 0.65 تعیین شد.

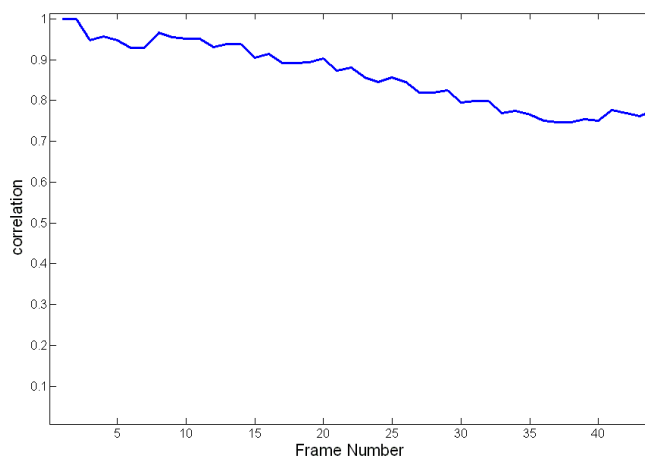
شکل ۶-۸، تعدادی از تصاویر مربوط به قسمتی از یک دنباله ویدئویی را نشان می‌دهد که شامل ۵۰ فریم می‌باشد. در این رشته سر یک حرکت انتقالی همزمان بسمت چپ و جلو دارد. همانطور که منحنی تغییرات مقدار محاسبه شده برای همبستگی متقابل نشان می‌دهد در فریم‌ها نهایی با نزدیک شدن زیاد سر به دوربین مقدار کرولیشن کاهش می‌یابد.



شکل ۶-۸ (الف) : پنجره‌های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



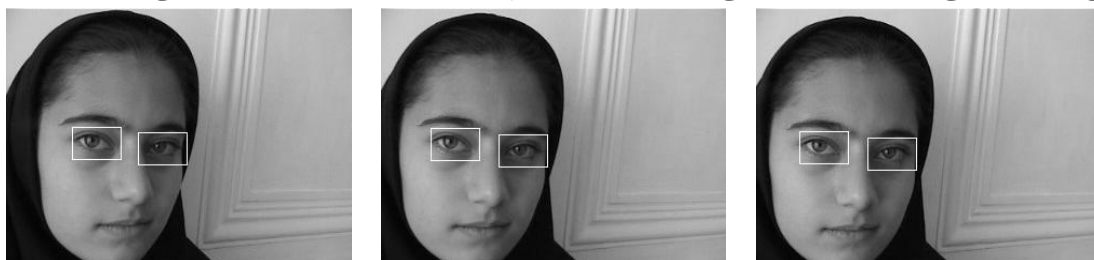
(ب) : منحنی تغییرات تابع همبستگی متقابل برای چشم راست



(ج) : منحنی تغییرات تابع همبستگی متقابل برای چشم چپ

شکل ۶-۸: نتایج الگوریتم ردیابی چشم در حرکت انتقالی سر

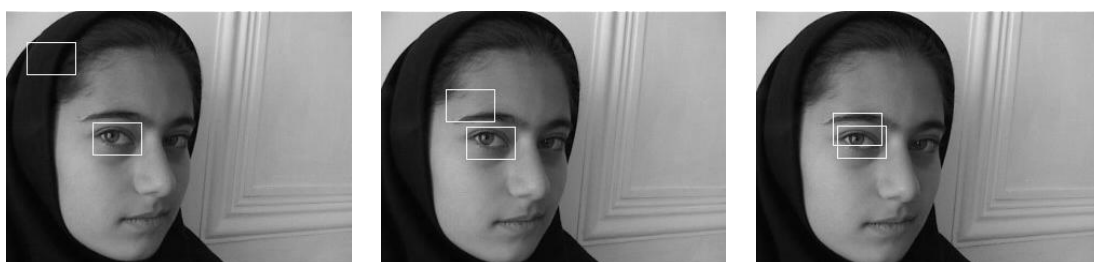
همانطور که گفته شد یکی از عواملی که باعث ایجاد خطا در الگوریتم می‌شود، چرخش سر بصورتی است که هر یک از چشم‌ها از تصویر ناپدید شوند در دنباله دیگر که تعدادی از فریم‌های آن در شکل ۶-۹ نشان داده شده است، سر یک حرکت انتقالی بسمت جلو داشته و سپس چرخشی عمقی بسمت چپ داشته است. با چرخش سر به میزان زیاد ردیابی چشم چپ با مشکل مواجه شده و در نهایت مکان چشم چپ گم می‌شود. منحنی تغییرات همبستگی متقابل برای چشم چپ نیز گویای این حقیقت می‌باشد.



(۳)

(۲)

(۱)

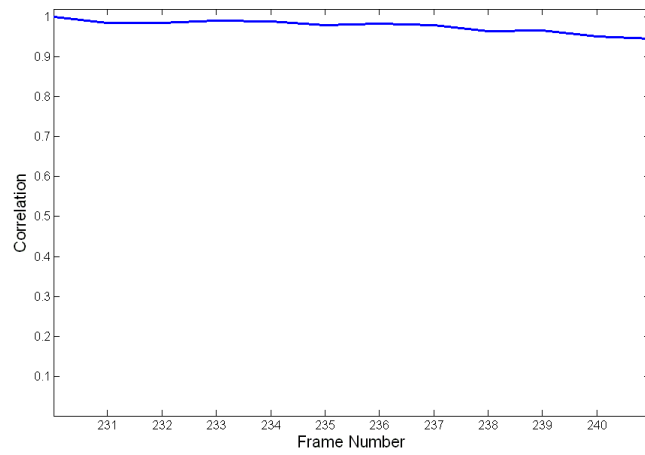


(۶)

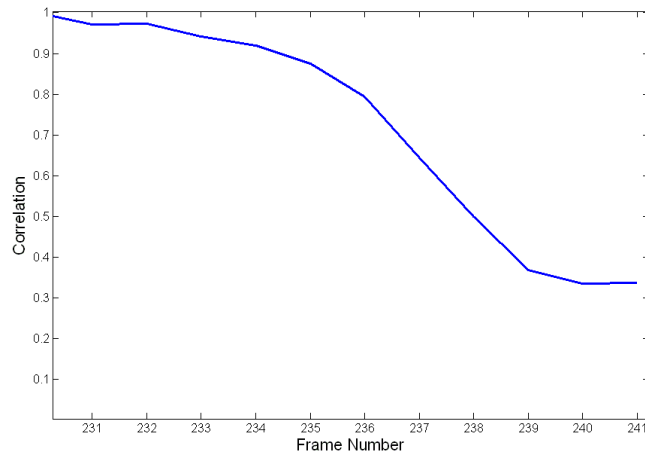
(۵)

(۴)

شکل ۶-۹ (الف) : پنجره‌های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



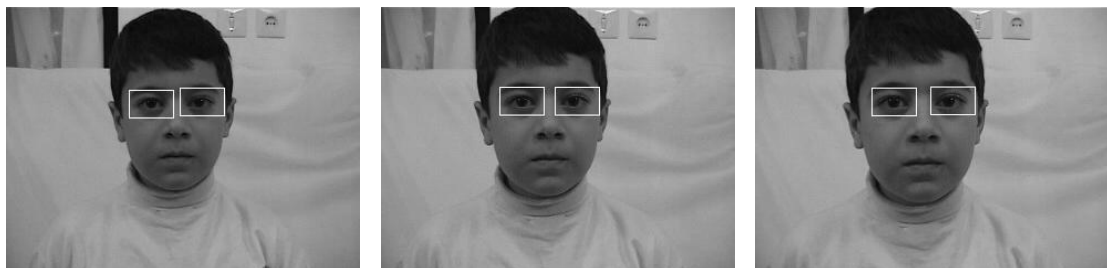
شکل ۹-۶ (ب) : منحنی تغییرات تابع همبستگی متقابل برای چشم راست



(ج) : منحنی تغییرات تابع همبستگی متقابل برای چشم چپ

شکل ۹-۶ : نتایج الگوریتم ردیابی چشم در چرخش عمقی سر

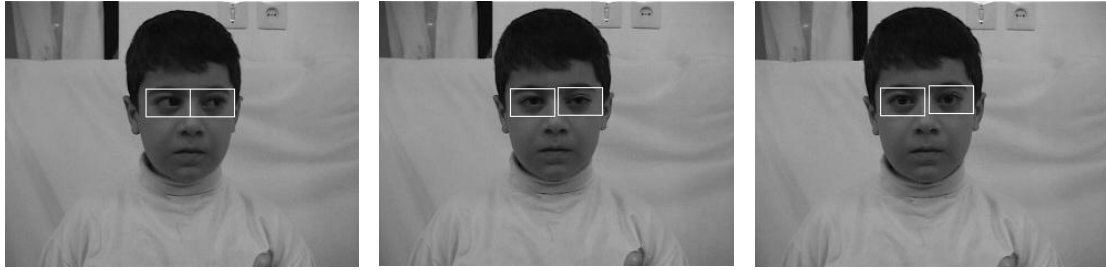
شکل ۱۰-۶ تصاویر مربوط به دنباله دیگری را نشان می دهد که بازهم چرخش سر در عمق باعث ایجاد خطا در ردیابی چشم چپ شده است. در این مرحله نیاز به تشخیص مجدد مکان چشم می باشد.



(۳)

(۲)

(۱)

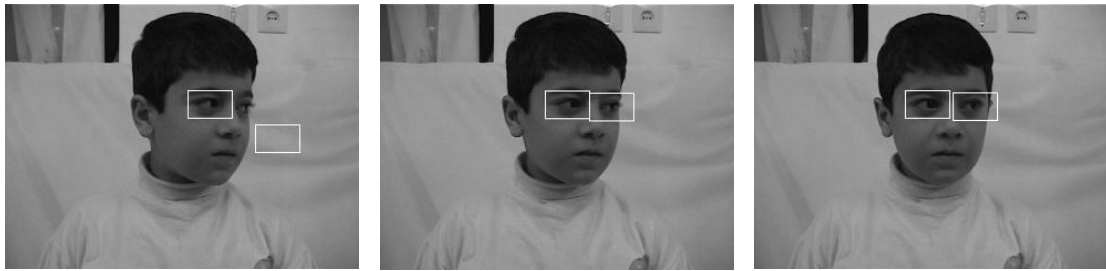


(۶)

(۵)

(۴)

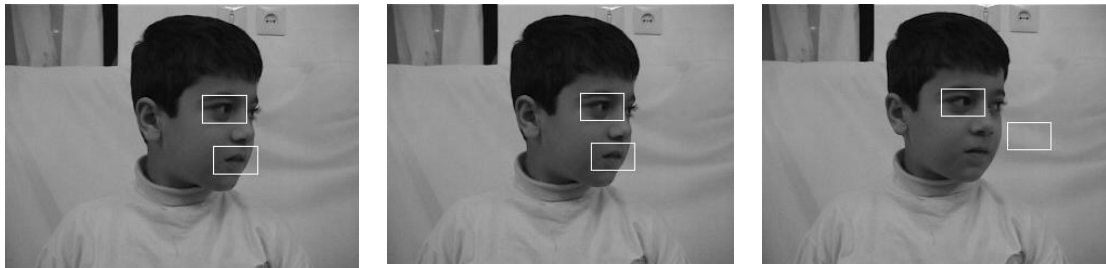
شکل ۶-۱۰ (الف): پنجره‌های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



(۹)

(۸)

(۷)

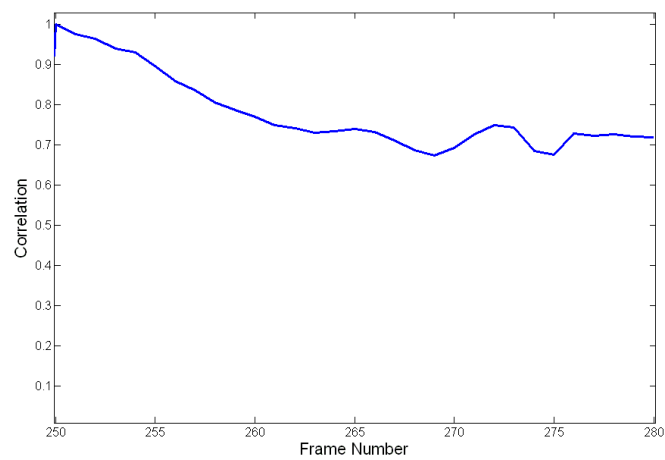


(۱۲)

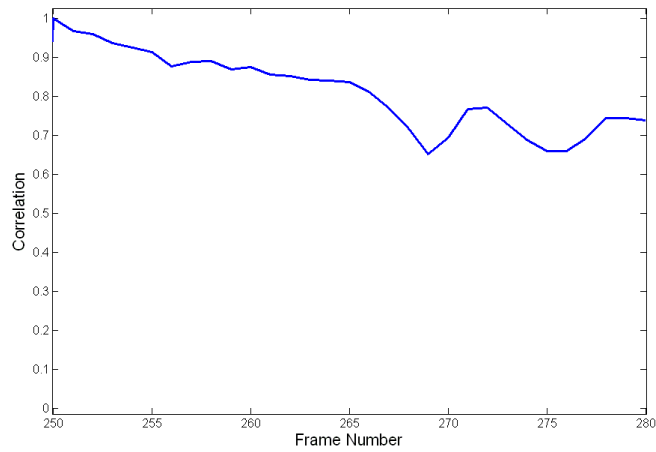
(۱۱)

(۱۰)

شکل ۶-۱۰ (الف): پنجره‌های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



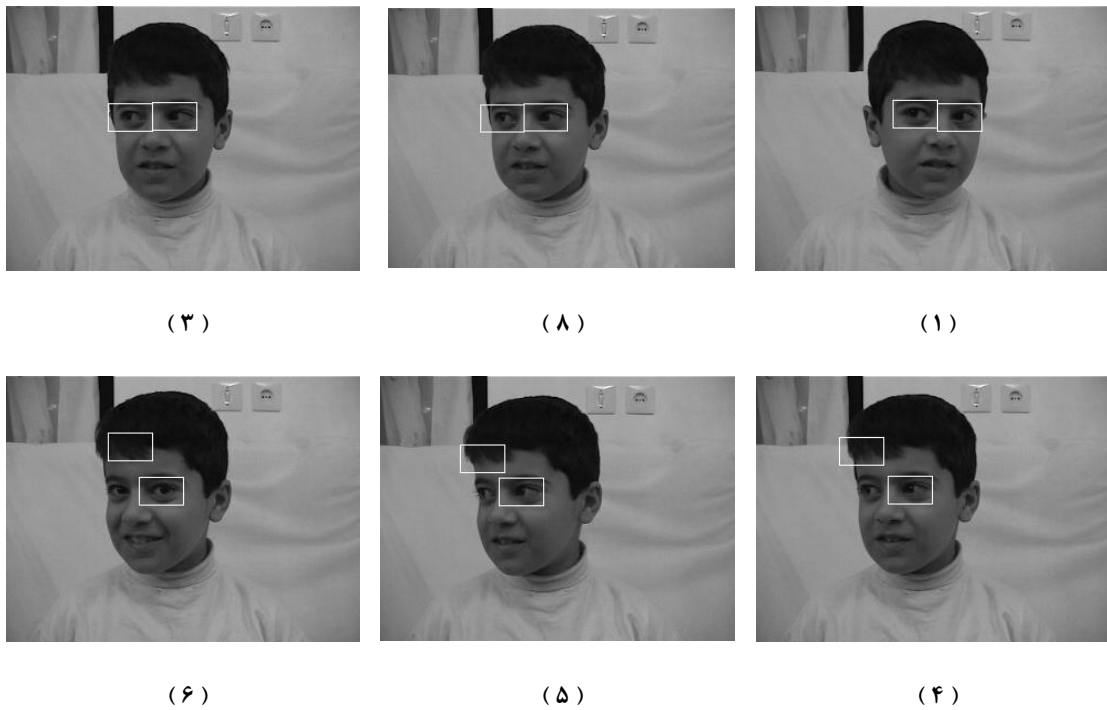
شکل ۶-۱۰ (ب): منحنی تغییرات تابع همبستگی متقابل برای چشم راست



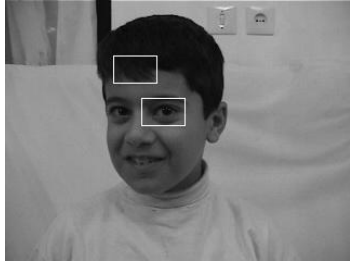
(ج) : منحنی تغییرات تابع همبستگی متقابل برای چشم چپ

شکل ۶-۱۰: نتایج الگوریتم ردیابی در حرکت انتقالی و چرخش عمقی سر

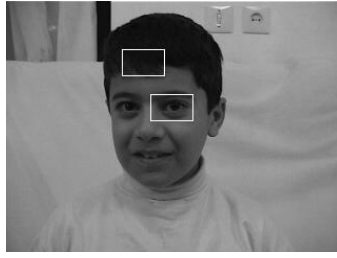
شکل ۶-۱۱ دنباله دیگری از چرخش عمقی سر بسمت راست را نشان می‌دهد. همانطور که در شکل دیده می‌شود با چرخش سر بسمت راست، ردیابی چشم راست با مشکل مواجه شده‌است.



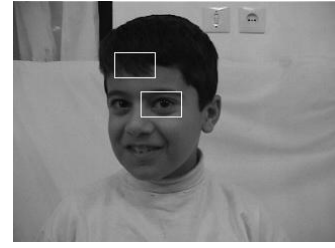
شکل ۶-۱۱ (الف) : پنجره‌های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



(۹)

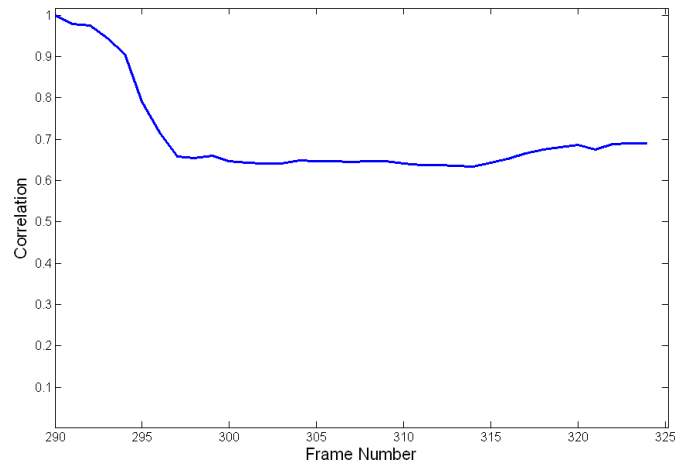


(۸)

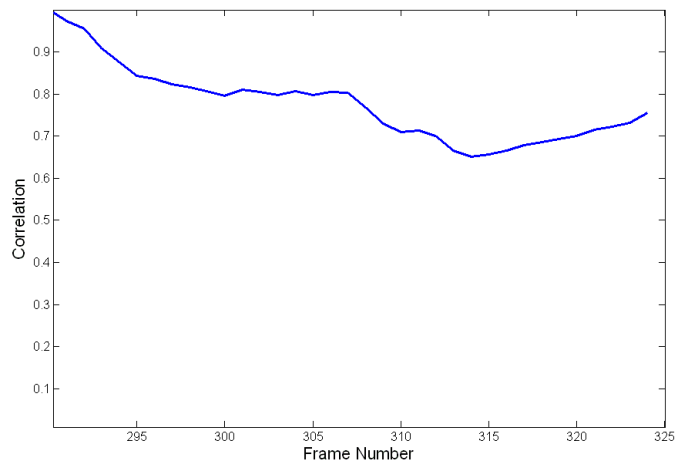


(۷)

(الف) : پنجره‌های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



(ب) : منحنی تغییرات تابع همبستگی متقابل برای چشم راست



(ج) : منحنی تغییرات تابع همبستگی متقابل برای چشم چپ

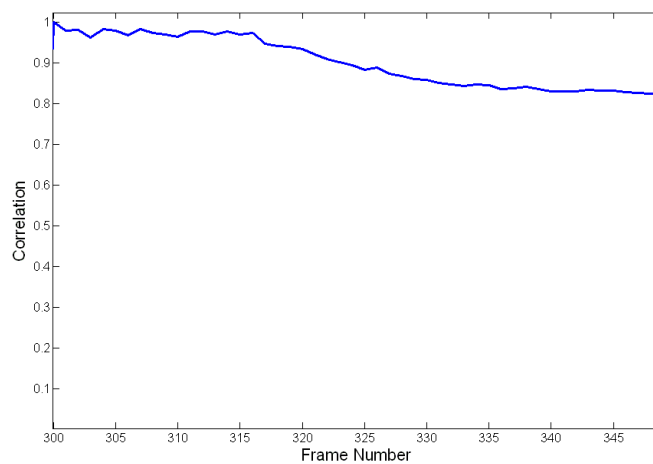
شکل ۶-۱۱: نتایج الگوریتم ردیابی در حرکت انتقالی و چرخش عمقی سر

۳۰ دنباله از پایگاه داده مورد استفاده در این بخش مربوط به افرادی بودند که از عینک استفاده می کردند. نتایج بدست آمده برای این افراد نشان می دهد الگوریتم در دنبال کردن چشمها با وجود عینک مقاوم می باشد. البته چنانچه فریم عینک در این افراد ضخیم باشد به نحویکه باعث پوشانده شدن ناحیه چشمها

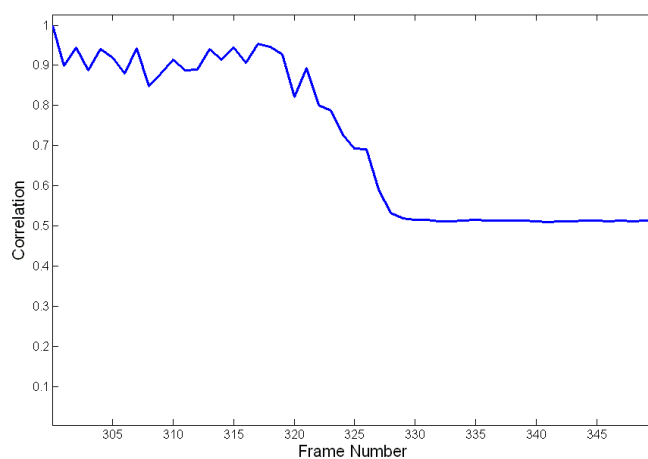
شود، الگوریتم دچار خطا می‌شود. نمونه‌ای از این دنباله‌ها در شکل ۶-۱۲ نشان داده شده است. در این دنباله چرخش عمقی سر باعث خطای الگوریتم شده است.



شکل ۶-۱۲ (الف): پنجره‌های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



(ب): منحنی تغییرات تابع همبستگی متقابل برای چشم راست



(ج) : منحنی تغییرات تابع همبستگی متقابل برای چشم چپ

شکل ۶-۱۲ : نتایج الگوریتم ردیابی در استفاده از عینک

همانطور که گفته شد بسته شدن چشم بطور کامل که در پلک زدن اتفاق می افتد یکی دیگر از عوامل ایجاد خطا و از دست دادن ناحیه چشمها در فرایند ردیابی می شود (شکل ۶-۱۳).



(۳)

(۲)

(۱)

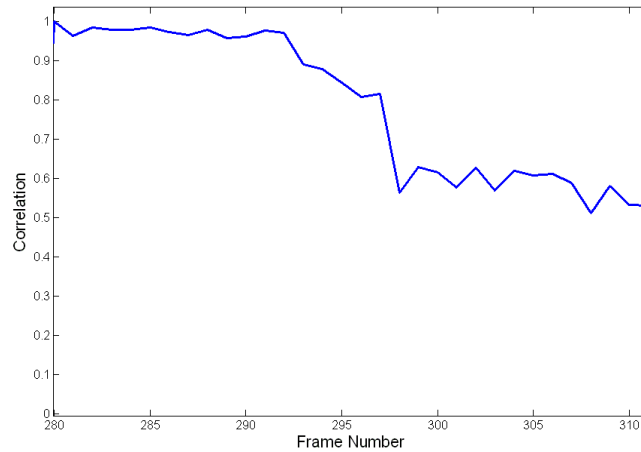


(۶)

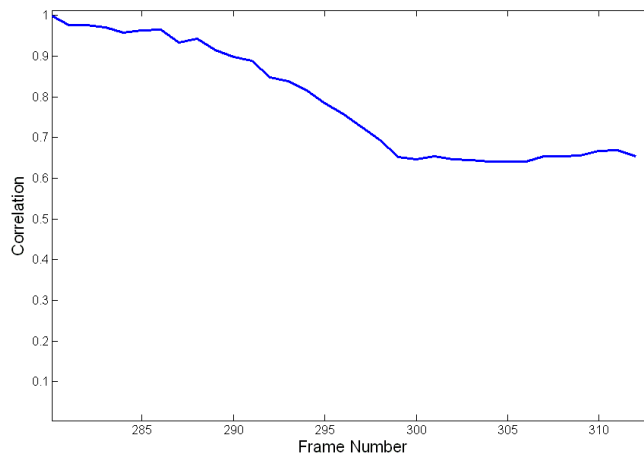
(۵)

(۴)

(الف) : پنجره های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



(ب) : منحنی تغییرات تابع همبستگی متقابل برای چشم چپ



(ج) : منحنی تغییرات تابع همبستگی متقابل برای چشم چپ

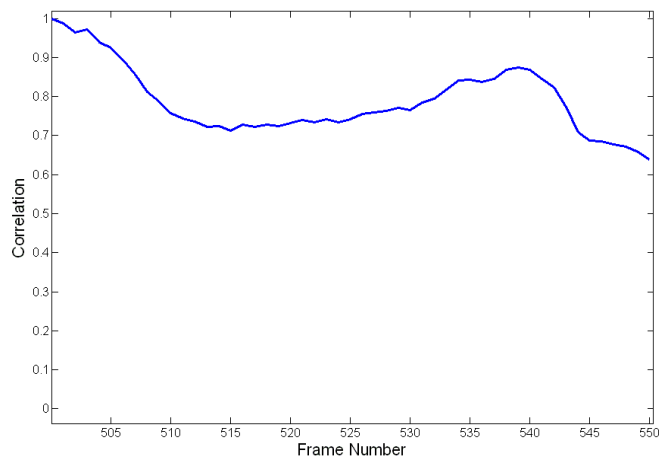
شکل ۶-۱۳: نتایج الگوریتم ردیابی در پلک زدن

یکی دیگر از حالاتی که امکان وقوع آن در حرکت سر زیاد است چرخش سطحی سر در مقابل دوربین می‌باشد. برای ارزیابی الگوریتم در صورت چرخش سر در صفحه در تمامی دنباله‌های تهیه شده، از افراد خواسته شد تا سر را در دو جهت ممکن در صفحه بچرخانند. پاسخ‌های بدست آمده نشان می‌دهد الگوریتم نسبت به چرخش سر در صفحه عمود بر محور آپتیکی دوربین مقاوم می‌باشد. در دنباله نشان داده شده در شکل ۶-۱۴ پلک زدن باعث خطا در نتیجه شده است.

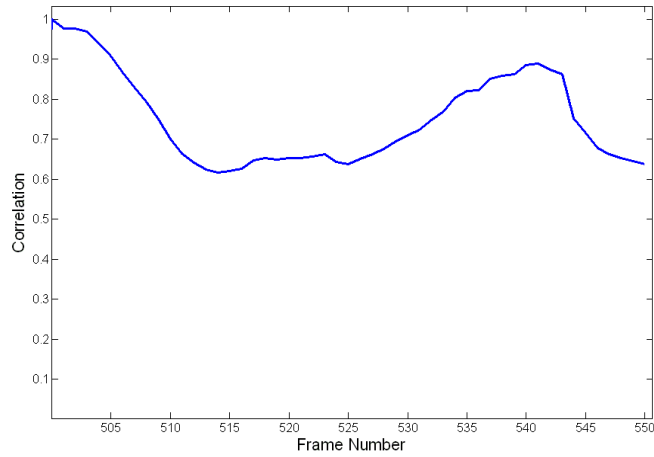




شکل ۶-۱۴ (الف) : پنجره‌های تعیین شده بعنوان چشم توسط الگوریتم پیشنهادی



(ب) : منحنی تغییرات تابع همبستگی متقابل برای چشم راست

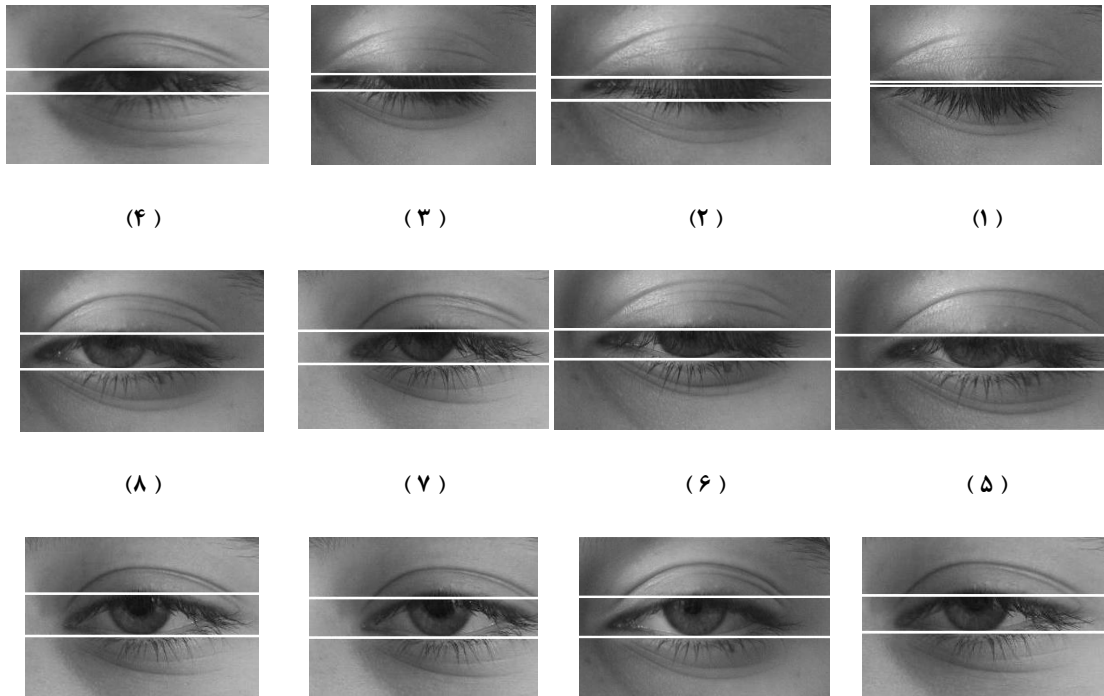


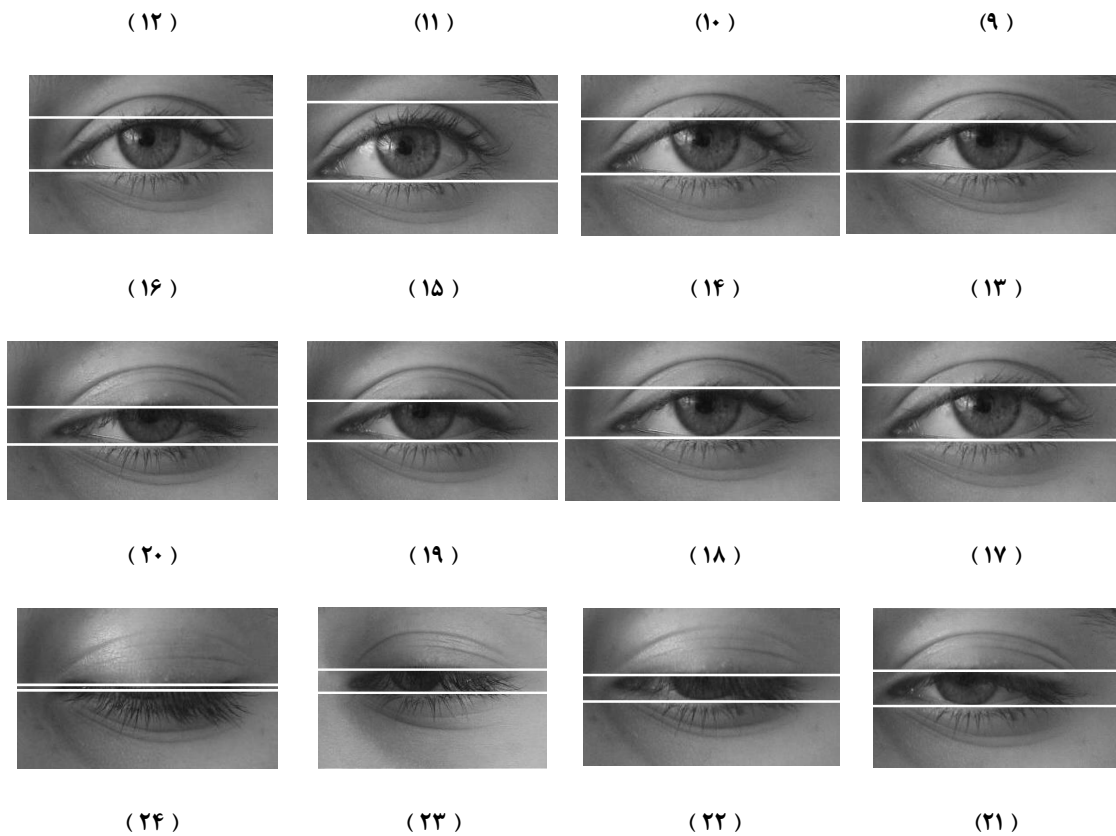
(ج) : منحنی تغییرات تابع همبستگی متقابل برای چشم چپ

شکل ۶-۱۴: نتایج الگوریتم ردیابی در حرکت سطحی سر در صفحه دوربین

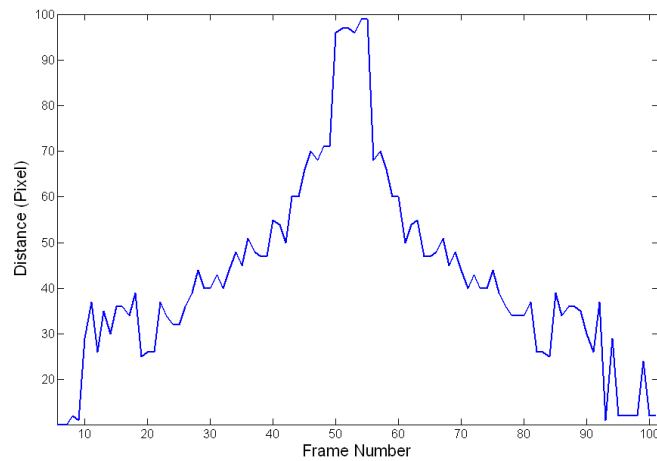
۶-۲-۳ - نتایج الگوریتم تعیین فاصله بین پلک‌ها

در این بخش نتایج حاصل از اعمال الگوریتم پیشنهادی برای تعیین فاصله بین پلک‌ها بر روی دنباله‌های ویدئویی تهیه شده ارائه می‌گردد. نتایج بدست آمده نشان می‌دهد در فریم‌هایی که چهره در آن بصورت تمام‌رخ قرار گرفته و ناحیه چشم‌ها بطور کامل مشخص است، الگوریتم کارایی بسیار قابل قبولی دارد. شکل ۶-۱۵ یک نمونه‌ای از نتایج بدست آمده بر روی یک دنباله ویدئویی را نشان می‌دهد. خطوط سفید رنگ بیانگر محل‌های تشخیص داده‌شده توسط الگوریتم بعنوان پلک بالا و پایین چشم هستند. این شکل فرایند پلک را برای یک شخص نشان می‌دهد. فاصله بین پلک‌ها برحسب شماره فریم بصورت یک منحنی در شکل ۶-۱۶ نشان داده شده‌است.



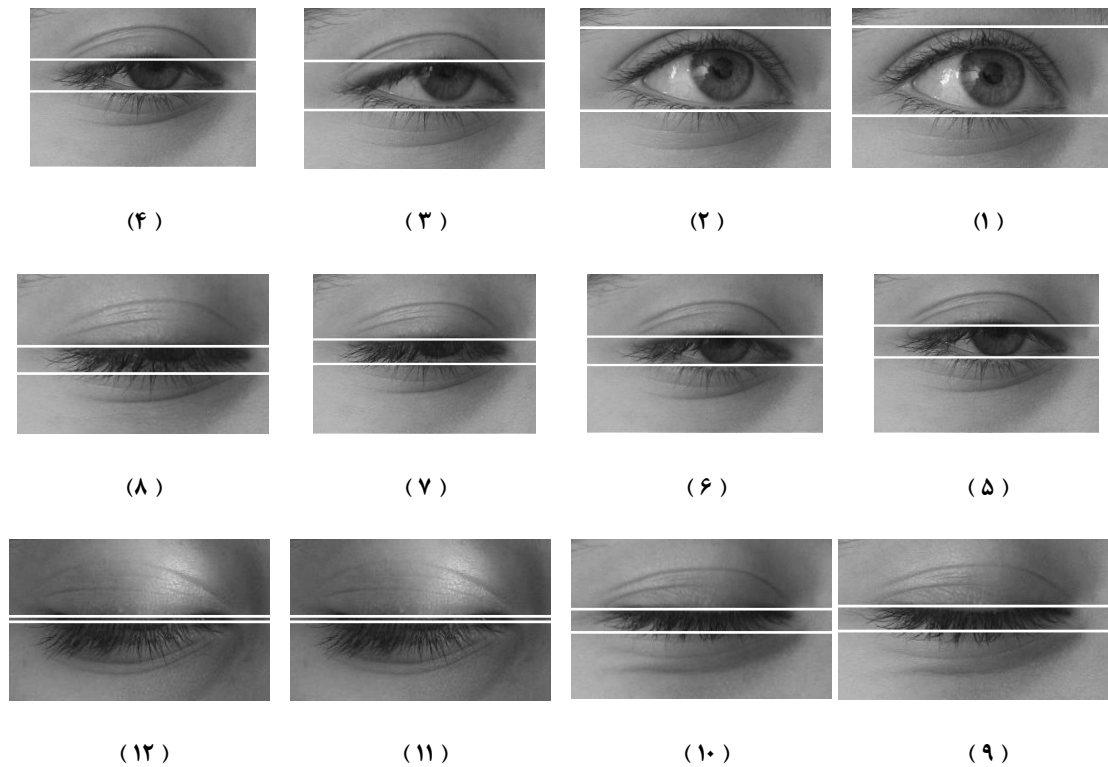


شکل ۶-۱۵: تعیین مکان پلک‌ها در پلک زدن

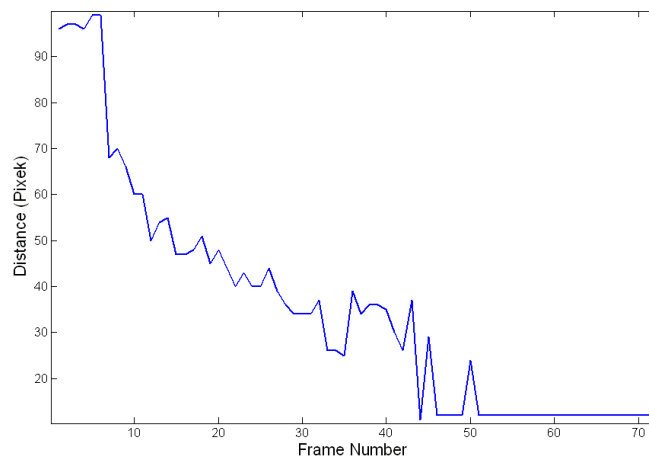


شکل ۶-۱۶: منحنی نمایش دهانه فاصله پلک‌ها در پلک زدن

در شکل ۶-۱۷ فاصله بین پلک‌ها برای محدوده زمانی دیگری مورد بررسی قرار گرفته است. همانطور که در تصاویر دیده می‌شود که پلک‌ها تغییر حالت آنی نداشته‌اند و پس از بسته شدن برای مدت طولانی به همان حالت باقی مانده‌اند. این دنباله مربوط به خواب آلودگی در فرد می‌باشد. شکل ۶-۱۸ منحنی تغییرات فاصله بین پلک‌ها را برای این دنباله نشان می‌دهد.

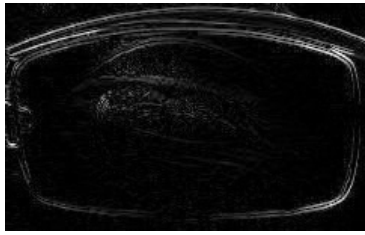


شکل ۶-۱۷: تعیین فاصله بین پلک‌ها در خواب آلودگی



شکل ۶-۱۸: منحنی نمایش دهانه فاصله پلک‌ها در خواب آلودگی

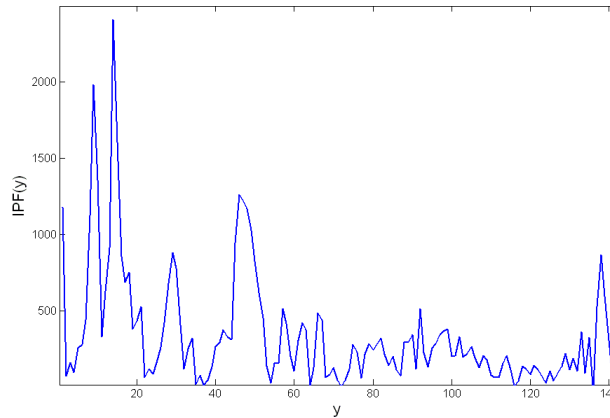
یکی از عواملی که تشخیص مکان پلک‌ها را با چالش مواجه می‌کند وجود عینک می‌باشد. نتایج نشان می‌دهد در صورتیکه فریم عینک ضخامت زیادی داشته باشد و بنحوی روی چشم قرار بگیرد که باعث پوشانده شدن پلک‌ها شود، الگوریتم دچا خطا شده و محل فریم عینک را بجای محل پلک در تصویر مشخص می‌کند. برای مقابله با این مشکل می‌توان ناحیه مورد بررسی در اطراف چشم را تا حد امکان کوچک در نظر گرفت. تا وجود فریم عینک در آن منتفی شود البته در صورت کوچک بودن فریم یا قرار گرفتن آن روی پلک، این راه‌حل مؤثر نخواهد بود. شکل ۶-۱۹ نمونه‌ای از این حالت را که پلک بالای چشم با فریم عینک پوشانده شده‌است، نشان می‌دهد.



شکل ۱۹-۶ (ب): تصویر اندازه گرادیان روشنایی در جهت عمودی



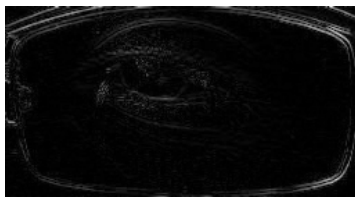
شکل ۱۹-۶ (الف): تصویر ناحیه چشم



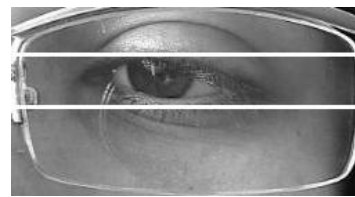
(ج) : منحنی تابع نگاشت جمع کننده در جهت افقی

شکل ۱۹-۶ : پاسخ الگوریتم در رابطه با تصاویر همراه با عینک

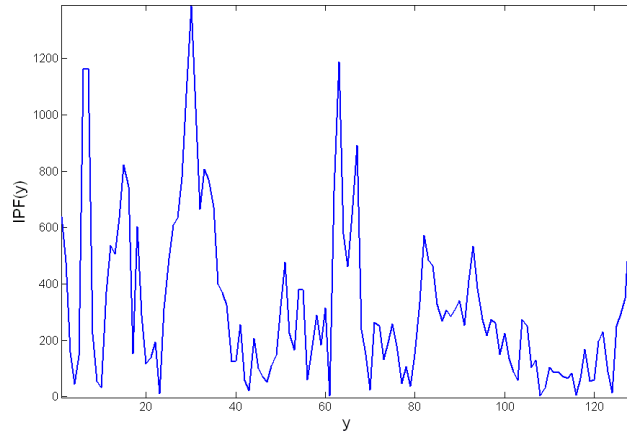
شکل ۲۰-۶ حالتی را نشان می دهد که تصویر ناحیه چشم، از مکانی نزدیکتر به محل پلکها جدا شده است. بدیهی است در صورت کوچک بودن فریم عینک یا بزرگی ابعاد چشم این روش با خطا مواجه می شود.



(ب): تصویر اندازه گرادیان روشنایی در جهت عمودی



(الف): تصویر ناحیه چشم



(ج) : منحنی تابع نگاشت جمع کننده در جهت افقی

شکل ۶-۲۰: پاسخ الگوریتم پس از تصحیح ناحیه چشم

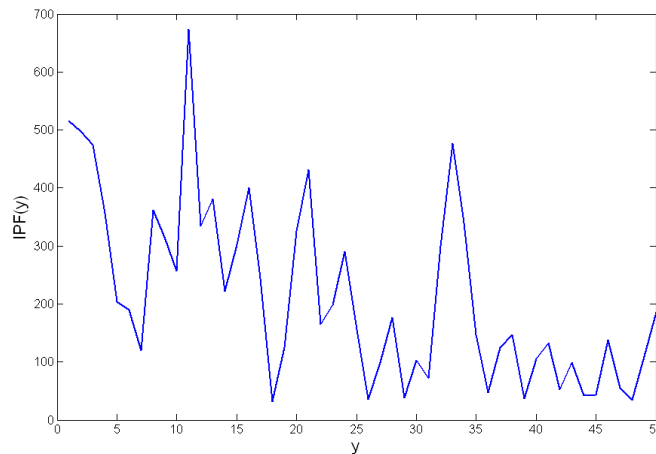
یکی دیگر از مسائلی که در زمان استفاده از عینک ایجاد می شود انعکاس نور در شیشه عینک می باشد. آزمایش بر روی دنباله های ویدئویی تهیه شده با وضعیت انعکاس در شیشه عینک، نشان می دهد که الگوریتم نسبت به این وضعیت مقاوم می باشد. دو مورد از فریم های مربوط به این شرایط به همراه گرادیان روشنایی تصویر در جهت عمودی و تابع نگاشت محاسبه شده روی این تصویر در شکل های ۶-۲۱ و ۶-۲۲ نشان داده شده است.



(ب): تصویر اندازه گرادیان روشنایی در جهت

(الف): تصویر ناحیه چشم

عمودی



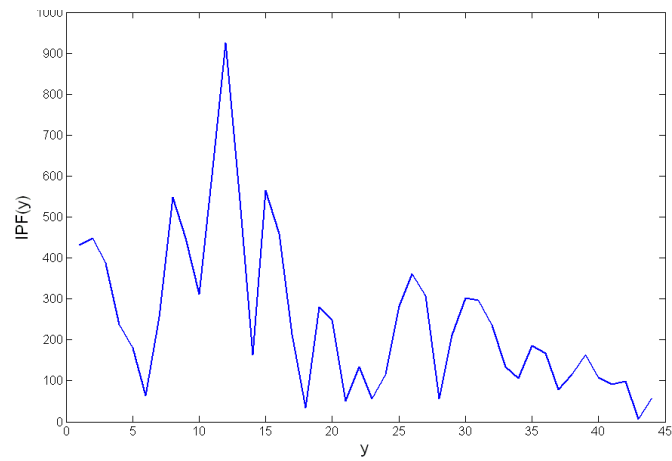
(ج) : منحنی تابع نگاشت جمع کننده در جهت افقی

شکل ۶-۲۱: پاسخ الگوریتم در رابطه با انعکاس نور در شیشه عینک



تصویر اندازه‌گرادیان روشنایی در جهت عمودی

(الف): تصویر ناحیه چشم



(ج) : منحنی تابع نگاشت جمع کننده در جهت افقی

شکل ۶-۲۲: پاسخ الگوریتم در رابطه با انعکاس نور در شیشه عینک

نتیجه گیری و پیشنهادات

در این پایان نامه موضوع تشخیص چشم و ردیابی آن در تصاویر دو بعدی مورد بررسی قرار گرفت. در بخش آخر مسأله تعیین فاصله بین پلک‌ها که کاربرد زیادی در سیستم‌های تعامل انسان و کامپیوتر دارد مطالعه شد. با پس از بررسی فعالیت‌های انجام شده تا کنون و تحلیل نتایج بدست آمده توسط آنها، الگوریتم‌های پیشنهادی برای بهبود روش‌های موجود ارائه گردید. روش ارائه شده در این تحقیق برای تشخیص چشم‌ها در دسته روش‌های مبتنی بر ظاهر قرار می‌گیرد. در این الگوریتم، ابتدا ساختار توپوگرافیک تصویر استخراج می‌شود و با تکیه بر منحصر بفرد بودن ساختار توپوگرافیک چشم، مکان چشم‌ها در تصاویر مشخص می‌شود. برای بهبود نتایج و مقاوم نمودن الگوریتم نسبت به تبدیل‌های هندسی شباهتی از ممان‌های تغییر ناپذیر استفاده شد. به این منظور در ابتدا تصویر برچسب خورده توپوگرافیکی به زیر تصویرهایی تجزیه شد که هر کدام شامل تنها یک برچسب بودند. سپس مقدار توابع پیشنهادی Hu [۶۵] که ترکیبی از ممان‌های تغییر ناپذیر می‌باشند، برای این زیر تصویرها محاسبه شد. مقادیر این توابع عناصر بردار ویژگی می‌باشند. در الگوریتم پیشنهادی بردار ویژگی دارای ۳۵ عنصر می‌باشد. به منظور آموزش کلاسه‌بند از سه گروه آموزشی شامل تصاویر چشم راست، چشم چپ و غیر چشم استفاده شد. با در نظر گرفتن توزیع نرمال برای ویژگی‌های استخراج شده، بردار میانگین و ماتریس کواریانس برای گروه‌های مذکور محاسبه شد. در مرحله آزمون با استفاده فاصله مالهانونوبیس بین نواحی کاندید و سه گروه موجود مکان چشم‌ها مشخص می‌گردد. استفاده از ممان‌های تغییر ناپذیر نتایج مفیدی از جمله کاهش ابعاد بردار ویژگی و افزایش سرعت در پی داشتند. در الگوریتم پیشنهادی برای ردیابی چشم‌ها از تطبیق الگو استفاده شده است. بدلیل اینکه پروسه تشخیص چشم‌ها، زمان بیشتری در مقایسه با تطبیق الگو نیاز دارد، در اولین فریم از دنباله ویدئویی مکان چشم‌ها مشخص می‌شود و سپس ناحیه چشم‌ها در فریم‌های بعدی ردیابی می‌گردد. در این الگوریتم از مقدار تابع همبستگی متقابل برای تعیین صحت نتایج استفاده می‌شود. در آخرین بخش از این تحقیق، الگوریتمی به منظور تعیین فاصله بین پلک‌ها ارائه شد. این الگوریتم با توجه به این نکته که تغییر سطح روشنایی در محل پلک‌ها زیاد می‌باشد، محل پلک‌ها را تعیین می‌کند. به منظور مقاوم ساختن الگوریتم نسبت به تغییر نورپردازی، از گرادیان تصویر استفاده می‌شود. پس از محاسبه دامنه گرادیان تصویر، از تابع نکاشت در جهت عمودی استفاده می‌شود. این تابع سطری از تصویر که گرادیان در آن بیشترین مقدار را دارد بعنوان مکان پلک‌ها در نظر می‌گیرد.

در بخش پایانی نیز نتایج حاصل از پیاده‌سازی الگوریتم‌های پیشنهادی از زوایای گوناگونی مورد بررسی قرار گرفت و نقاط قوت و ضعف آن بیان شد. تحلیل ما در این مورد به ارائه راه‌حل‌های گوناگونی جهت بهبود نتایج منجر گردید. راه‌حل‌های پیشنهاد شده به شرح ذیل می‌باشند:

- کاهش حساسیت الگوریتم تشخیص چشم نسبت به تغییر پارامترهای فیلترهای

موجود در این الگوریتم: یکی از مهمترین مشکلات موجود در الگوریتم تشخیص چشم،

حساسیت الگوریتم نسبت به تغییر پارامترهای فیلتر هموارساز گوسی و تغییر ابعاد پنجره

مربوط به فیلتر مشتق‌گیر می‌باشد. راه‌حلی که برای این مشکل پیشنهاد می‌گردد استفاده از پردازش چند رزولوشنی در الگوریتم تشخیص چشم‌ها می‌باشد. کاملاً قابل پیش‌بینی است که این عمل پیچیدگی الگوریتم را افزایش داده و باعث کاهش سرعت سیستم می‌شود.

- **بهبود بردار ویژگی:** بهبود عملکرد الگوریتم تشخیص چشم توسط هرچه غنی‌تر نمودن بردار ویژگی همواره مورد توجه بوده‌است. استفاده از روش‌های انتخاب ویژگی‌ها به منظور استفاده از ویژگی‌های برتر در بردار ویژگی، می‌تواند عملکرد الگوریتم را به میزان زیادی بهبود بخشد.

- **بهبود الگوریتم ردیابی چشم‌ها در جهت کاهش خطای منفی کاذب:** همانطور که نتایج الگوریتم ردیابی چشم‌ها نشان داد این الگوریتم در مواردیکه سر در صفحه دوربین می‌چرخد، خطای منفی کاذب ایجاد می‌کند. یعنی با وجود ردیابی چشم‌ها بطور صحیح، پاسخ الگوریتم در دسته پاسخ‌های خطا قرار می‌گیرد. به نظر می‌رسد محاسبه تابع همبستگی متقابل بین دو ناحیه انتخاب شده بعنوان چشم‌ها توسط الگوریتم می‌تواند معیار مناسبی برای جلوگیری از خطا در این موارد باشد. راه‌حل دیگری که پیشنهاد می‌شود تصحیح الگوها در هر فریم می‌باشد بنحویکه الگوها در هر فریم در جهت خط واصل مردمک‌ها از چهره جدا شوند و در فریم بعدی از این الگوها برای تطبیق الگو استفاده شود.

- **بهبود روش تشخیص مکان پلک‌ها در نور ضعیف:** اکثر سیستم‌های تشخیص خواب در رانندگان که اخیراً پیشنهاد شده‌اند از تصویر برداری مادون قرمز استفاده می‌کنند. دلیل این موضوع، سرعت عملکرد بالا و پاسخ مناسب این الگوریتم در نور بسیار ضعیف می‌باشد. اما این نوع تصویر برداری نیاز به تجهیزات مختلف و گاهاً گران‌قیمتی دارد. روش پیشنهادی برای تعیین فاصله بین پلک‌ها دارای پیچیدگی محاسباتی پایین و سرعت بالایی می‌باشد، همچنین بدلیل استفاده از گرادیان به میزان زیادی نسبت به نورپردازی مقاوم است. اما پاسخ‌های الگوریتم پیشنهادی در نور بسیار ضعیف خطای زیادی دارد. به نظر می‌رسد استفاده از برجسب‌های توپوگرافیکی ناحیه چشم‌ها و برجسب چاله در مردمک که نسبت به نورپردازی بسیار مقاوم می‌باشند کمک شایانی به حل این مشکل کند.

ضمیمه الف

الف-۱ - فرم‌های جبری و تغییر ناپذیر با زمان^۱

رابطه چند جمله ای همگن دو متغیره (رابطه الف-۱) که بر حسب u و v است، فرم جبری و ی با به بیان ساده تر فرم باینری ممان از مرتبه p نامیده می شود [۴۸].

$$f = a_{p0}u^p + \binom{p}{1}a_{p-1,1}u^{p-1}v + \binom{p}{2}a_{p-2,2}u^{p-2}v^2 + \dots + \binom{p}{p-1}a_{1,p-1}u v^{p-1} + a_{0p}v^p \quad \text{الف-۱}$$

با استفاده از نمادسازی توسط Caley می توان نوشت:

$$f \equiv (a_{p0}; a_{p-1,1}; \dots; a_{1,p-1}; a_{0p})(u, v)^p. \quad \text{الف-۱}$$

جمله ای همگن $I(a)$ متشکل از الفرایب a_{p0}, \dots, a_{0p} ثابت جبری با وزن w نامیده می شود اگر:

$$I(a'_{p0}, \dots, a'_{0p}) = \Delta^w I(a_{p0}, \dots, a_{0p}),$$

که a'_{p0}, \dots, a'_{0p} الفرایب جدیدی است که با جای‌گزینی تبدیلی خطی عمومی زیر در شکل اصلی (رابطه الف-۱) حاصل می شود:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix}, \quad \Delta = \begin{vmatrix} \alpha & \gamma \\ \beta & \delta \end{vmatrix} \neq 0. \quad \text{الف-۲}$$

اگر $w = 0$ ثابت $I(a)$ را ثابت مطلق^۲ می نامیم، در غیر اینصورت آن را، ثابت وابسته می نامند [۵۸]. ثابت تعریف شده در بالا ممکن است به الفرایب بی‌ش از یک فرم بستگی داشته باشد. تحت تبدیلی خطی ویژه ای که در بخش الف-۲ بیان خواهد شد، Δ ممکن است محدود به دترمینان فوق نباشد. با حذف Δ در بین دو ثابت وابسته همواره یک متغیر مطلق جمع ناپذیر حاصل می گردد.

در مطالعه ثابتها، یک جفت x و y دی‌گر که تبدیلی آنها با توجه به رابطه الف-۲ به صورت زیر است محاسبه می گردد:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad \text{الف-۳}$$

تبدیلی بیان شده در رابطه الف-الف-۳ تبدیلی cogredient نامیده می شود. متغیرهای x و y مقادیر covariant نام دارند و u و v مقادیر contravariant نام دارند. این پارامترها رابطه ثابت را برآورده می سازند.

¹ Algebraic Forms and Invariant

² Absolute Invariant

الف-۴

$$ux + vy = u'x' + v'y'$$

مطالعه ثابتهای جبری توسط Boole, Cayley, Sylvester آغاز شد و قویاً توسط سالی ری-ن ادامه پی-دا کرد. ممانهای آماری که در بخش الف-۲ مطرح خواهند شد از روی نتایج ثابتهای جبری منتج می شوند.

الف-۲ - ممانهای تغییر ناپذیر^۱

الف-۲-۱ - تئوری پایه برای ممانهای تغییر ناپذیر^۲

تابع مولد ممان با عامل نمایی قابل گسترش به صورت سری زیر تعریف می گردد:

$$M(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{p=0}^{\infty} \frac{1}{p!} (ux + vy)^p \rho(x, y) dx dy. \quad \text{الف-۵}$$

با جابه جا کردن اپراتورهای خطی انتگرال و سی-گما خواهیم داشت:

$$M(u, v) = \sum_{p=0}^{\infty} \frac{1}{p!} (\mu_{p_0}, \dots, \mu_{0_p})(u, v)^p. \quad \text{الف-۶}$$

با تبدیلی رابطه الف-۲ به رابطه الف-۳ و علامت گذاری ضرایب x' و y' به عوامل تبدیلی یافته $(ux + vy)$ با u' و v' به ترتیب و y به y' به طور مشابه با اعمال هردو رابطه (الف-۱) و (الف-۲) به (الف-الف-۵) خواهیم داشت:

$$M_1(u', v') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{p=0}^{\infty} \frac{1}{p!} (u'x' + v'y')^p \rho'(x', y') \frac{1}{|J|} dx' dy' \quad \text{الف-۷}$$

که $|J|$ دترمینان ماتریس ژاکوبین تبدیلی (الف-۲) و $\rho'(x', y') \equiv \rho(x, y)$ تابع

تولید ممان بعد از تبدیلی است. اگر ممان دوم تبدیلی یافته μ'_{pq} به صورت زیر تعریف گردد:

$$\mu'_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{p=0}^{\infty} \frac{1}{p!} (x')^p (y')^q \rho'(x', y') dx' dy', \quad p, q = 0, 1, 2, \dots \quad \text{الف-۸}$$

سپس خواهیم داشت:

$$M_1(u', v') = \frac{1}{|J|} \sum_{p=0}^{\infty} \frac{1}{p!} (\mu'_{p_0}, \dots, \mu'_{0_p})(u', v')^p. \quad \text{الف-۹}$$

¹ Moment Invariants

2 A Fundamental Theorem of Moment Invariants

در تئوری ثابتهای جبری، مرسوم است که قانون تبدیلی برای الفرایب a در فرم جبری رابطه ممان مرکزی ($\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q \rho(x, y) dx dy$) مشابه $x^{p-r} y^r$ در رابطه زیر است:

$$(ux + vy)^p = (x^p, x^{p-1}y, \dots, y^p)(u, v)^p. \quad \text{الف-۱۰}$$

از روابط (الف-۵، الف-۴، الف-۵، الف-۶ و الف-۷) کاملاً روشن است که روابط مشابهی بین مرتبه p ام به جز ضریب $|J|^{-1}$ برقرار است. بنابراین تئوری اساسی زیر شکل می‌گیرد.

تئوری اساسی ممان‌های آماری: اگر فرم جبری از درجه p دارای ثابت جبری زیر باشد:

$$I(a'_{p0}, \dots, a'_{0p}) = \Delta^w I(a_{p0}, \dots, a_{0p}), \quad \text{الف-۱۱}$$

ممان از مرتبه p ثابتی کاملاً مشابه با یک عامل اضافی $|J|$ دارد:

$$I(\mu'_{p0}, \dots, \mu'_{0p}) = |J| \Delta^w I(\mu_{p0}, \dots, \mu_{0p}). \quad \text{الف-۱۲}$$

الف-۲-۲ - تشابه ممان‌های ثابت^۱

تحت تبدیلی تشابه، یعنی تغییری در ابعاد، هر ضریب از هر فرم جبری ثابت خواهد بود:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \alpha \rightarrow const. \quad \text{الف-۱۳}$$

$$a'_{pq} = \alpha^{p+q} a_{pq},$$

که در این رابطه α ، دترمینان نخواهد بود. برای ثابت ممان داریم:

$$\mu'_{pq} = \alpha^{p+q+2} \mu_{pq}. \quad \text{الف-۱۴}$$

با حذف α بین رابطه مرتبه صفرام:

$$\mu' = \alpha^2 \mu$$

و "یک"های باقی مانده، ثابت مطلق زیر بدست خواهد آمد:

$$\frac{\mu'_{pq}}{(\mu')^{(p+q)/2+1}} = \frac{\mu_{pq}}{(\mu)^{(p+q)/2+1}}, \quad p+q = 2, 3, \dots \quad \text{and} \quad \mu'_{10} = \mu'_{01} \equiv 0 \quad \text{الف-۱۵}$$

الف-۲-۳ - ممان‌های ثابت متعامد^۲

تحت تبدیلی چرخش متعامد زیر:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{الف-۱۶}$$

داریم:

$$J = \begin{vmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{vmatrix} = +1. \quad \text{الف-۱۷}$$

بنابراین ثابتهای ممان دقیقا برابر ثابتهای جبری خواهند شد. اگر به ممانها مثل ضرایب فرم جبری نگاه کنیم:

$$(\mu_{p0}, \dots, \mu_{0p})(u, v)^p \quad \text{الف-۱۸}$$

تحت تبدیلی Contragredient زیری:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix}, (\mu_{p0}, \dots, \mu_{0p})(u, v)^p \quad \text{الف-۱۹}$$

ثابتهای ممان با روش جبری زیر بدست می آید:

$$\begin{bmatrix} U \\ V \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}, \quad \begin{bmatrix} U' \\ V' \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix}, \quad \text{الف-۲۰}$$

سپس تبدیلی متعامد به رابطه ساده زیر تبدیلی می گردد

$$U' = Ue^{-i\theta}, \quad V' = Ve^{i\theta}. \quad \text{الف-۲۱}$$

با جایگزینی روابط الف-الف-۲۱ و الف-الف-۲۱ در الف-۱۷ رابطه زیر را خواهیم داشت:

$$\begin{aligned} (I_{p0}, \dots, I_{0p})(U, V)^p &\equiv (\mu_{p0}, \dots, \mu_{0p})(u, v)^p \equiv (\mu'_{p0}, \dots, \mu'_{0p})(u', v')^p \\ &\equiv (I'_{p0}, \dots, I'_{0p})(Ue^{-i\theta}, Ve^{i\theta})^p, \end{aligned} \quad \text{الف-۲۲}$$

که I_{p0}, \dots, I_{0p} و I'_{p0}, \dots, I'_{0p} ضرایب متناظر بعد از جایگزینی هستند. از رابطه بالا در U و V ضرایب $U^{p-1}V$ در دو طرف رابطه می بایست برابر باشند. بنابراین:

$$\begin{aligned} I'_{p0} &= e^{ip\theta} I_{p0}; \quad I'_{p-1,1} = e^{i(p-2)\theta} I_{p-1,1}; \dots; \\ I'_{1,p-1} &= e^{i(p-2)\theta} I_{1,p-1}; \quad I'_{0p} = e^{-ip\theta} I_{0p}. \end{aligned} \quad \text{الف-۲۳}$$

(p+1) ثابت ممان مستقل تحت تبدیلی صحیح متعامد وجود دارد و $\Delta = e^{i\theta}$ که چیزی متفاوت از دترمینان تبدیلی است.

از روی دو عبارت اول رابطه الف-۲۱ مشاهده می شود $I_{r,p-r}$ مزدوج مختلط $I_{p-r,r}$ است و

$$I_{p0} = \mu_{p0} - i \binom{p}{1} \mu_{p-1,1} - \binom{p}{2} \mu_{p-2,2} + i \binom{p}{3} \mu_{p-3,3} + \dots + (-i)^p \mu_{0p},$$

$$I_{p-1,1} = (\mu_{p0} + \mu_{p-2,2}) - i(p-2)(\mu_{p-1,1} + \mu_{p-3,3}) + \dots + (-i)^{p-2}(\mu_{2,p-2} + \mu_{0p}),$$

$$I_{p-2,2} = (\mu_{p0} + 2\mu_{p-2,2} + \mu_{p-4,4}) - i(p-4)(\mu_{p-1,1} + \mu_{p-3,3} + \mu_{p-5,5})$$

$$+ \dots + (-i)^{p-4}(\mu_{4,p-4} + 2\mu_{2,p-2} + \mu_{0p}),$$

الف-۲۴

$$I_{p-r,r} = [(\mu_{p0}; \mu_{p-2,2}; \dots; \mu_{p-2r,2r})(1,1)^r; (\mu_{p-1,1}; \mu_{p-3,3}; \dots; \mu_{p-2r-1,2r+1})(1,1)^r; \dots;$$

$$(\mu_{2r,p-2r}; \mu_{2r+2,p-2r-2}; \dots; \mu_{0p})(1,1)^r](1,-i)^{p-2r}, p-2r > 0.$$

$$I_{p/2,p/2} = \mu_{p0} + \binom{p/2}{1} \mu_{p-2,2} + \binom{p/2}{2} \mu_{p-4,4} + \dots + \mu_{0p}, p = \text{even}.$$

لازم به ذکر است که این $I(p+1)$ ها به توابع خطی مستقل از خطی از μ ها هستند و برعکس.

برای تبدیلی نامناسب متعامد زیری یعنی چرخش و انعکاس:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \begin{vmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{vmatrix} = -1.$$

الف-۲۵

بطور مشابه :

$$U' = V e^{i\theta} \quad V' = U e^{-i\theta}$$

الف-۲۶

$$I'_{p0} = e^{-ip\theta} I_{0p}; \quad I'_{p-1,1} = e^{-i(p-2)\theta} I_{1,p-1}; \dots; \quad I'_{1,p-1} = e^{i(p-2)\theta} I_{p-1,1}; \quad I'_{0p} = e^{ip\theta} I_{p0},$$

الف-۲۷

ثابتهای متعامد اولی ن بار توسط Boole مطالعه شد و الگوریتم فوق توسط Sylvester ارائه گردید.

الف-۲-۴ - A Complete System of Absolute Orthogonal Moment Invariant

از روی روابط الف-الف-۲۳ و الف-۲۳ سیستم ممانهای تغییر ناپذیر زیری با حذف عامل $e^{i\theta}$ حاصل می گردد:

برای ممانهای درجه دوم، دو تابع تغییر ناپذیر عبارتند از:

$$I_{11}, \quad I_{20}, I_{02}.$$

الف-۲۸

برای ممان درجه سوم، سه تابع تغییر ناپذیر عبارتند از:

$$I_{30} I_{03}, \quad I_{21} I_{12}, \quad (I_{30} I_{12}^3 + I_{03} I_{21}^3).$$

الف-۲۹

همچنین چهارمین ممان وابسته به ممانهای مرتبه سوم عبارتست از:

الف-۳۰

$$\frac{1}{i}(I_{30}I_{12}^3 - I_{03}I_{21}^3)$$

رابطه جبری بین چهار ممان داده شده در روابط الف-۲۹ و الف-الف-۲۹ وجود دارد. سه تای اول که در (الف-۲۸) آمده، توابع تغییر ناپذیر و آخرین آنها که در رابطه الف-الف-۳۰ بیان شده است فقط تحت چرخش و تغییری علامت ثابت است. این ممان از نوع ثابت نسبت به کجی^۱ نامگذاری می گردد. یکی تابع تغییر ناپذیر دیگر از روی ممانهای دوم و سوم به ترتیب زیر قابل ساخت است:

الف-۳۱

$$(I_{20}I_{12}^2 + I_{02}I_{21}^2).$$

برای ممانهای از مرتبه $p \geq 4$ ، $[p/2]$ انتگرال $p/2$ نامیده شده و ثابتهای زیر را خواهیم داشت:

الف-۳۲

$$I_{p0}I_{0p}, I_{p-1,1}I_{1,p-1}; \dots; I_{p-r,r}I_{r,p-r}; \dots$$

اگر p فرد باشد همچنین خواهیم داشت:

الف-۳۳

$$I_{p/2}I_{p/2}.$$

و با ترکیب با $(p-2)$ امین ممان $[p/2-1]$ ثابت خواهیم داشت.

الف-۳۴

$$(I_{p-1,1}I_{0,p-2} + I_{1,p-1}I_{p-2,0}), (I_{p-2,2}I_{1,p-3} + I_{2,p-2}I_{p-3,1})$$

$$(I_{p-r,r}I_{r-1,p-r+1} + I_{r,p-r}I_{p-r+1,r-1}), p - 2r > 0$$

در صورت ترکیب با ممان مرتبه دوم، اگر p فرد باشد داریم:

الف-۳۵

$$(I_{[p/2],[p/2]+1}^2 I_{20} + I_{[p/2]+1,[p/2]}^2 I_{02}),$$

اگر p زوج باشد:

الف-۳۶

$$(I_{p/2-1,p/2+1}I_{20} + I_{p/2+1,p/2-1}I_{02}),$$

بنابراین ما همواره $(p+1)$ تابع تغییر ناپذیر مطلق مستقل خواهیم داشت. با تغییری مجموع بالا به تفاضل نیز توابع تغییر ناپذیر با کجی بدست خواهند آمد.

همه توابع مستقل ممانها باهم یکی سیستم کامل را شکل می دهند، یعنی برای هر تابع داده شده می توان آنرا برحسب توابع مستقل فوق گسترش داد.

¹ Skew invariant

الف-۲-۵- ممان‌های تغییر ناپذیر تحت تبدیل خطی^۱

بر اساس تئوری ثابت‌های جبری تحت تبدیلات خطی عمومی (رابطه الف-الف-۱۶) مشهود است که عامل Δ دترمینان تبدیلی است. برای تبدیلات خطی J نیز دترمینان است. برای سادگی فرض می‌کنیم A, B, C و a, b, c, d بی‌انگرممانهای دوم و سوم هستند، سپس ما می‌توانیم دو فرم باینری را بر حسب این ممانها بنویسیم:

$$(A, B, C)(u, v)^2 \quad (a, b, c, d)(u, v)^3. \quad \text{الف-۳۷}$$

بر اساس تئوری ثابت‌های جبری، چهار ثابت مستقل جبری بصورت زیر خواهند بود:

$$\begin{aligned} I_1 &\equiv AC - B^2, \\ I_2 &\equiv (ad - bc)^2 - 4(ac - b^2)(bd - c^2), \\ I_3 &\equiv a^2C^3 - 6abBC^2 + 6acC(2B^2 - AC) + ad(6ABC - 8B^3) \\ &\quad + 9b^2AC^2 - 18bcABC + 6bdA(2B^2 - AC) \\ &\quad + 9c^2A^2C - 6cdBA^2 + d^2A^3, \end{aligned} \quad \text{الف-۳۸}$$

که در آن $w=2,4,6$. برای ممان مرتبه صفر داریم:

$$\mu' = |J| \mu. \quad \text{الف-۳۹}$$

با دانستن اینکه $\Delta^2 = |J|^2$ چهار ثابت ممان مطلق زیر قابل حصول خواهند بود.

$$\frac{I_1}{\mu^3}, \quad \frac{I_2}{\mu^7}, \quad \frac{I_3}{\mu^5}, \quad \frac{I_4}{\mu^7}. \quad \text{الف-۴۰}$$

همچنین ثابت دی‌گری با نام ثابت کجی (skew invariant) I_5 با وزن ۹ که به ممانهای A, B, C و a, b, c, d بستگی دارد نیز وجود دارد. این رابطه را همچنان می‌توان به صورت زیر نرمالیزه کرد:

$$\frac{\Delta}{|J|} \left(\frac{I_5}{\mu^{10}} \right), \quad \text{الف-۴۱}$$

که $\Delta/|J|$ علامت دترمینان را تعیین می‌کند.

با شمارش تعداد روابط در بین ممانها و تعداد پارامترهای استفاده شده، نشان داده می‌شود که چهار بزرگترین عدد ثابت مستقل ممکن برای این مورد است. روشهای بسیاری برای پیدا کردن ثابتهای جبری ارائه شده‌اند.

¹ Moment Invariants under General Linear Transformation

الف-۳- توصیف و شناسایی الگوها

هر الگوی تصویری مانند تصویر اجزای چهره و چشم و یا تصویر حروف الفبا و ارقام همواره با یک تابع توزیع $\rho(x, y)$ قابل نمایش هستند. در این رابطه x و y محورهای در جهت افقی و عمودی تصویر می‌باشند و ارزش p در پیکسل به موقعیت (x, y) معرف سطح روشنایی (در تصاویر سطح خاکستری) و یا رنگ (در تصاویر رنگی) در آن می‌باشد. آشکار است که این الگو را می‌توان با مجموعه ممان‌های دوبعدی آن $\{m_{pq}\}$ توصیف کرد. با در نظر گرفتن ارتباط بین ممانهای مرکزی و ممانهای معمولی با داشتن یکی، دیگری بدست خواهد آمد. در صورتی که این ممانهای مرکزی با استفاده از ممانهای آماری مشابه هنجار گردند مجموعه‌ای از ممانها برای مشخص کردن الگوی خاص قابل حصول است و یک ممان به تنهایی توانایی توصیف الگوی مورد نظر را ندارد. آشکار است که این پارامترها مستقل از مکان و ابعاد الگوی مورد نظر هستند.

در ادامه روش محاسبه ممانهای مستقل از چرخش بیان می‌شود. در حقیقت از نظر تئوری بی‌نیازی ممان با در نظر گرفتن محورهای اصلی وجود دارد. اما واضح است که در عمل تنها تعداد محدودی از آنها قابل استفاده هستند. در حقیقت تنها تعداد اندکی از این ممانها برای بسیاری از کاربردها کافی هستند.

الف-۳-۱- روش محورهای اصلی^۱

در روابط الف-۲۲ و الف-۲۳ فرض کنید $p = 2$ در این صورت ثابتهای ممان z را خواهیم داشت:

$$\begin{aligned}(\mu'_{20} - \mu'_{02}) - i 2\mu'_{11} &= e^{j2\theta} [(\mu_{20} - \mu_{02}) - i 2\mu_{11}], \\ (\mu'_{20} - \mu'_{02}) + i 2\mu'_{11} &= e^{-j2\theta} [(\mu_{20} - \mu_{02}) + i 2\mu_{11}], \\ \mu'_{20} + \mu'_{02} &= \mu_{20} + \mu_{02}.\end{aligned}$$

الف-۴۲

اگر زاویه θ از روی رابطه اول در رابطه الف-۴۱ طوری محاسبه گردد که $\mu'_{11} = 0$ در این صورت خواهیم داشت:

$$\tan 2\theta = \frac{+2\mu_{11}}{\mu_{20} - \mu_{02}}$$

الف-۴۳

محورهای x', y' با استفاده از θ ای که در رابطه الف-۴۲ صدق کند محاسبه شده و محورهای اساسی نامیده می‌شوند. با اضافه کردن محدودیتهای نظری $\mu'_{20} > \mu'_{02}$ و $\mu'_{30} > 0$ ، به طوری که تعیین می‌شود. ممانهایی که برای چینی محورهای اساسی تعیین شده اند مستقل از چرخش خواهند بود.

در صورت استفاده از ممانهای بالاتر قابلیت جدایی نمونه‌ها افزایش می‌یابد. اگر ثابتهای داده شده در روابط الف-۲۳ و الف-۲۴ استفاده گردند، ممانهای بالاتر مرتبط با محورهای اساسی به آسانی بدست می‌آیند. این روابط به شکل دیگری نیز مفید خواهند بود. به عنوان نمونه برای $p = 3$ داریم:

¹ The Method of Principal Axes

$$(\mu'_{30} - 3\mu'_{12}) - i(3\mu'_{21} - 3\mu'_{03}) = e^{i3\theta}[(\mu_{30} - 3\mu_{12}) - i(3\mu_{21} - 3\mu_{03})],$$

$$(\mu'_{30} + \mu'_{12}) - i(\mu'_{21} + \mu'_{03}) = e^{i\theta}[(\mu_{30} + \mu_{12}) - i(\mu_{21} + \mu_{03})].$$

الف-۴۴

دو رابطه باقی مانده که مزدوج مختلط این دو هستند در اینجا ذکر نشده اند. اگر θ و ممان چهارم معلوم باشند، ممان های مشابه به آسانی از روی روابط بالا بدست می آیند.

فهرست مراجع

1. J. Reulen, J.T. Marcus, D. Koops, F. de Vries, G. Tiesinga, K. Boshuizen, J. Bos, Precise recording of eye movement: the iris technique, part 1, *Med. Biol. Eng. Comput.* 26 (1) (1988) 20–26.
 2. S. C. Kim, M. Sked, and Q. Ji, Non-intrusive Eye Gaze Tracking Under Natural Head Movements, *Proceedings of the 26th Annual International Conference of the IEEE EMBS, San Francisco, CA, USA*, September 1-5, 2004.
 3. R. Chellappa, C.L. Wilson, S. Sirohey, Human and machine recognition of faces: a survey, *Proc. IEEE* 83 (5) (1995) 705-740.
 4. B. Fasel, J. Luetin, Automatic facial expression analysis: a survey, *Pattern Recognition* 36 (2003) 259-275.
 5. L. Gan, B. Cui, W. Wang, Driver Fatigue Detection Based on Eye Tracking, *Proceedings of the 6th World Congress on Intelligent Control and Automation, June 21 - 23, 2006.*
 6. G.C. Feng, P. C. Yuen, Variance projection function and its application to eye detection for human face recognition. *Pattern Recognition Letters, Volume 19, Issue 9 (July 1998)*, 1998, pp 899 – 906
 7. H. Gu, G. Su, C. Du. Feature Points Extraction from Faces, *Image and Vision Computing*, 26-28 November, 2003, New Zealand; available online at http://sprg.massey.ac.nz/ivcnz/Proceedings/IVCNZ_28.pdf.
 8. Z. Zhou, X. Geng, Projection Functions for Eye Detection, *Pattern Recognition* 37, pp:1049-1056, 2004.
 9. L. Tao, H. K. Kwan, Automatic Localization of Human eyes in complex background, *IEEE*, pp: 669-672, 2002.
۱۰. ج. نصیری، س. خانچی، ح. پوررضا، آشکارسازی چشم در تصاویر رنگی، چهارمین کنفرانس ماشین بینایی و پردازش تصویر ایران، ۱۳۸۵.
۱۱. ف. شاهجویی، آشکارسازی چشم‌ها و دهان در تصویرهای دیجیتالی، چهارمین کنفرانس ماشین بینایی و پردازش تصویر ایران، ۱۳۸۵.
12. M. Pardas, Extraction and Tracking of the Eyelids. *International Conference on Acoustics, Speech and Signal Processing ICASSP 2000*, 4: 2357-2360, Istanbul, Turkey, June 2000; available online at http://gpstsc.upc.es/imatge/pub/ps/ICASSP00_pardas.pdf.
 13. T.D'Orazio, M. Leo, G. Cicirelli, A. Distante, An Algorithm for real time eye detection in face image, *Proceeding of the 17th International Conference on Pattern Recognition (ICPR'04)*, IEEE, 2000.
 14. H. Tan, Y. j. Zang, R. Li, Robust Eye Extraction Using Deformable Template and Feature Tracking Ability, *ICICS-PCM, 15-18 December 2003.*
 15. C. Tomasi and T. Kanade, "Detection and tracking of point features," *Technical Report. CMU-CS-91-132*, Carnegie Mellon University, Pittsburgh, PA, April 1991.
 16. L. Jin, X. Yaun, S. Satoh, J. Li, L. Xia, A hybrid classifier for precise and robust eye detection. *The 18th International Conference on Pattern Recognition (ICPR'06)*, 2006.
 17. Z. Niu, Sh. Shan, Sh. Yan, X. Chen, W. Gao, 2D Cascaded AdaBoost for eye Localization, *IEEE*, 2006.
 18. Z. Xingming, Z. Huangyuan, An Illumination Independent Eye Detection Algorithm, *Proceeding of the 18th International Conference on Pattern Recognition (ICPR'06)*, 2006.
 19. Jaepil KO, Eunju Kim and Hyeran Byun. A simple illumination algorithm for face recognition, *Proc. of PRICAI*, 2002.

20. T. Kawaguchi, M. Rizon, Iris detection using intensity and edge information, *Pattern Recognition* 36, 549-563, 2003.
21. C. H. Lin, J.L. Wu, Automatic facial feature extraction by genetic algorithms, *IEEE Trans. Image Process*, 834-845, 1999.
22. S. A. Suandi, T. S. Tai, S. Enokida, T. Ejima, *Horizontal Human Face Pose Determination Using Pupils and Skin Region Positions*, Springer-Verlag Berlin Heidelberg 2007, D. Mery and L. Rueda (Eds.) : PSIVT 2007, LNCS 4872, pp. 413–426, 2007.
23. Y. Tian, T. Kanade, and J. F. Cohn, Eye-state Action Unit Detection by Gabor Wavelets. In *Proceedings of International Conference on Multi-modal Interfaces (ICMI 2000)*, October, 2000; available online at <http://citeseer.ist.psu.edu/637068.html>
24. S. Sirohey, A. Rosenfeld, Z. Duric, A method of detecting and tracking irises and eyelids in video , *Pattern Recognition*, Vol.35 (2002), pp.1389–1401; availableonline at <http://cs.gmu.edu/~zduric/WebPages/Papers/PR-2002-sirohey.pdf>
25. M. J. Black, Y. Yacoob, Recognizing Facial Expressions in Image Sequences Using Local Parameterized Models of Image Motion. *International Journal of Computer Vision*. Vol. 25(1), October 1997, pp. 23 – 48; available online at <http://citeseer.ist.psu.edu/black97recognizing.html>
26. T. Moriyama, J. Xiao, J. F. Cohn, T. Kanade. Meticulously Detailed Eye Model and Its Application to Analysis of Facial Image. *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, 2004, pp. 629 – 634; available online at http://www.ri.cmu.edu/pubs/pub_4811.html
27. M. Pardas, Extraction and Tracking of the Eyelids. *International Conference on Acoustics, Speech and Signal Processing ICASSP 2000*,4: 2357-2360, Istanbul,Turkey, June 2000; available online at http://gpstsc.upc.es/imatge/pub/ps/ICASSP00_pardas.pdf
28. A. Garsia, F. M. Sanchez, A. Perez, J. L. Pedraza, R. Mendez, M. L. Cordoba, M. L. Munoz, A Distributed Real Time Eye-Gaze Tracking System, *IEEE*, 2003.
29. A. Fathi, M. T. Manzuri, Eye Detection and Tracking in Video Streams, *International Symposium on Communications and Information Technologies 2004 (ISCIT 2004)*, Sapporo, Japan, October 26-29, 2004
30. A.Haro, M.Flickner, I.Essa, Detecting and Tracking Eyes By Using Their Physiological Properties, *Dynamics and Appearance*, *IEEE*, 2000.
31. J. B. Hiley, A. H. Redekopp, R. F. Rezai, A Low Cost Human Computer Interface based on Eye Tracking, *Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA, Aug 30-Sept 3*,3226-3229, 2006.
32. D. W. Hansen, R. I. Hammoud, An improved likelihood model for eye tracking, *Computer Vision and Image Understanding* 106, 220-230, 2007.
33. P. Viola, M. Joones, Robust real-time face detection, *Proceedings of the International Conference on Computer Vision*, Vol. II, 2001.
34. E. Y. Kim, S. K. Kang, *Eye Tracking Using Neural Network and Mean-Shift*, Springer –Verlag Berlin Heidelberg, M. Gavrilova et al. (Eds.) : ICCSA 2006 LNCS 3982, pp. 1200-120, 2006.
35. R. S. Fris, T. E. de Campos, R. Marcondes Cesar Junior, *Detection and Tracking of Facial Features in Video Sequences*, Springer –Verlag Berlin Heidelberg, O.Cairo, L.E. Sucar, and F.J. Cantu (Eds.) : MICAI 2000, LNAI 1793, pp. 127-135, 2000.

36. Z. Zhu, Q. Ji, Robust real-time eye detection and tracking under variable lighting conditions and various face orientation, Elsevier Inc, Computer vision and Image Understanding 98 (2005) 124-154, 2005.
37. P. Jafari Moghadam Fard, M. H. Moradi, S. Parvaneh, Eye Tracking Using A Novel Approach, IFMBE Processong vol.14/4, pp.2407-2410.
38. G. Marcone, G. Martinelli, L. Lancetti, Eye Tracking in Image Sequences by Competitive Neural Networks, Neural Processing Letters 7: 133–138, 1998.
39. A. akhbarian azar, A. Akhbardeh, HUMAN COMPUTER INTERACTION-APPLYING FUZZY C-MEANS, RECURRENT NEURAL NETWORK AND WAVELET TRANSFORMS FOR VOLUNTARY EYE BLINK DETECTION, EUSIPCo, 2007.
40. Tian, Y.L., Kanade, T., Cohn, J., 2000. Dual-state parametric eye tracking. In: Proc. Internat. Conf. on Face and Gesture Recognition, vol. 2000, pp. 26–30, 2006.
41. Lucas, B., Kanade, T., 1981. An interactive image registration technique with an application in stereo vision. In: The 7th Internat. Joint Conf. on Arti.cial Intelligence, pp. 674–679.
42. H. Tan, Y. J. Zhang, Detecting eye blink states by tracking iris and eyelids, Pattern Recognition Letters 27 (2006) 667–675, 2006.
43. L. T. Watson, T. J. Laffey, and R. M. Haralick, "Topographic classification of digital image intensity surfaces using generalized splines and the discrete cosine transformation," Comput. Vision Graphics Image Processing, vol. 29, pp. 143-167, Jan. 1985.
44. R. M. Haralick. L. T. Watson, and T. J. Laffey, "The topographic primal sketch," Int. J. Robotics Res., vol. 2, pp. 50-72, 1983.
45. L. Wang, T. Pavlidis, Direct Gray-Scale Extraction of Features for Character Recognition, IEEE TRANSACTIONS ON PA'ITERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 15, NO. 10, pp:1053-1067, OCTOBER 1993.
46. P. Meer and I. Weiss, "Smoothed differentiation filters for images," Cent. Automat. Res., Univ. of Maryland, 1989.
47. J. Wang, L. Yin, Detecting and Tracking Eyes Through Dynamic Terrain Feature Matching, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005.
48. M. Hu, Visual Pattern Recognition by Moment Invariants, IRE TRANSACTIONS ON INFORMATION THEORY, pp:179-187, 1962.
49. W. Pitts, W. S. McCulloch, How to know universals, Bull. Math. Biophys., vol. 9, pp. 127-147; September, 1947.
50. L. G. Roberts, Pattern recognition with an adaptive network,1960 IRE INTERNATIONAL CONVENTION RECORD, pt. 2, pp. 66-70, 1960.
51. Minsky, op. cit., pp. 11-12.
52. Jiang, X. Y. and Bunke, H., \Simple and Fast Computation of Moments", Pattern Recognition, Vol. 24, No.8, pp. 801-806, 1991.
53. Dudani, S. A., Breeding, K. J. and McGhee, R. B., \Aircraft Identi_cation by Moment Invariants", IEEE trans. Computing, Vol. 26, No. 1, pp. 39-45.
54. R. C. Gonzalez and R. E. Woods, Digital Image Processing (third edition), Reading, Massachusetts: Addison-Wesley, 1992.
55. D. I. Barnea, H. F. Silverman, A class of algorithms for fast digital image registration, IEEE Trans. Computers, 21, pp. 179-186, 1972.

56. J. P. Lewis, Fast Normalized Cross-Correlation, *Vision Interface*, p. 120-123,1995.
57. Y. Adachi, M. Asada, Correlation-Based Visual Tracking enhanced by Affine Motion Description.
58. SCHWEITZER, H., BELL, J. W., AND WU, F. 2002. Very fast template matching. In *European Conference on Computer Vision (ECCV)*. 358–372.
59. FIEGUTH, P. AND TERZOPOULOS, D. 1997. Color-based tracking of heads and other mobile objects at video frame rates. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 21–27.
60. COMANICIU, D., RAMESH, V., AND MEER, P. 2003. Kernel-based object tracking. *IEEE Trans. Patt. Analy. Mach.Intell.* 25, 564–575.
61. COMANICIU, D. 2002. Bayesian kernel tracking. In *Annual Conference of the German Society for Pattern Recognition*. 438–445.
62. JEPSON, A., FLEET, D., AND ELMARAGHI, T. 2003. Robust online appearance models for visual tracking. *IEEE Trans. Patt. Analy. Mach. Intell.* 25, 10, 1296–1311.
63. ITU-T, *sl Video Codec Test Model, TMN5, ITUTSS LBC-95, Study Group 15, Working Party 15/1, Expert's Group on Very Low Bitrate Visual Telephony*, Telenor Research, January, <http://web.fou.telenor.no/fou/DVC/tmn5/tmn5.html>, 1995.
64. L.D. Stefano, E. Viarani, *Vehicle Detection and Tracking Using the Block Matching Algorithm*.
65. B. Thorslund. *Electrooculogram Analysis and Development of a System for Defining Stages of Drowsiness*. Linköping University, Linköping, 2003.

Abstract

Eye detection and eye tracking techniques plays important roles in developing Human Computer Interaction (HCI) systems. Disabled aid applications, Driver fatigue detection and human identification are some examples. Efficiency of such systems is evaluated with respect to their accuracy in detecting and tracking of eyes in a video stream. The most important problems in detecting and tracking of eyes are head position i.e. rotation and orientation, obscure eyes due to glasses, hear and severe illumination conditions, etc. Therefore, many researches are being conducted to increase the reliability of these systems.

In the first part of this thesis we are concerned with developing a robust method for detection of eyes in face image. The proposed method provides a solution which is invariant against similarity transformation and also robust to severe illumination conditions. Instead of directly using intensity information for describing the eyes in the face image, we represent the image using topographic labels. Using a set of functions invariant to similarity transformation, we describe the labeled topographic image. Finally based on the extracted features a Bayesian classifier is designed to detect eyes in the face image.

In second part of the thesis, an eye tracking technique for video sequences is introduced. In order to track the eyes in the next frame, we search for a region at the neighborhood of each eye in current frame for which the correlation function is maximized. Thanks to the correlation function, the computational complexity of the method is considerably low. In the last part of research, a blink detection method is proposed, we simply measure the distance between eyelids for this purpose by mapping of gradient image of eyes in vertical direction.

The detection results applied on "XM2VTS" database confirms the superior performance in different imaging conditions. The result of eye tracking using the proposed method was very promising.



EYE DETECTION AND TRACKING IN 2D IMAGES

Thesis

Submitted in Partial Fulfillment of the
Requirements for the Degree of Master of Science (M.Sc.)
in Electrical Engineering

Department of Electrical Engineering
Shahrood University of Technology

By:

Saideh Ferdowsi

Supervisor:

Alireza Ahmadyfard

Advisor:

Ali Soleymani

February 2007