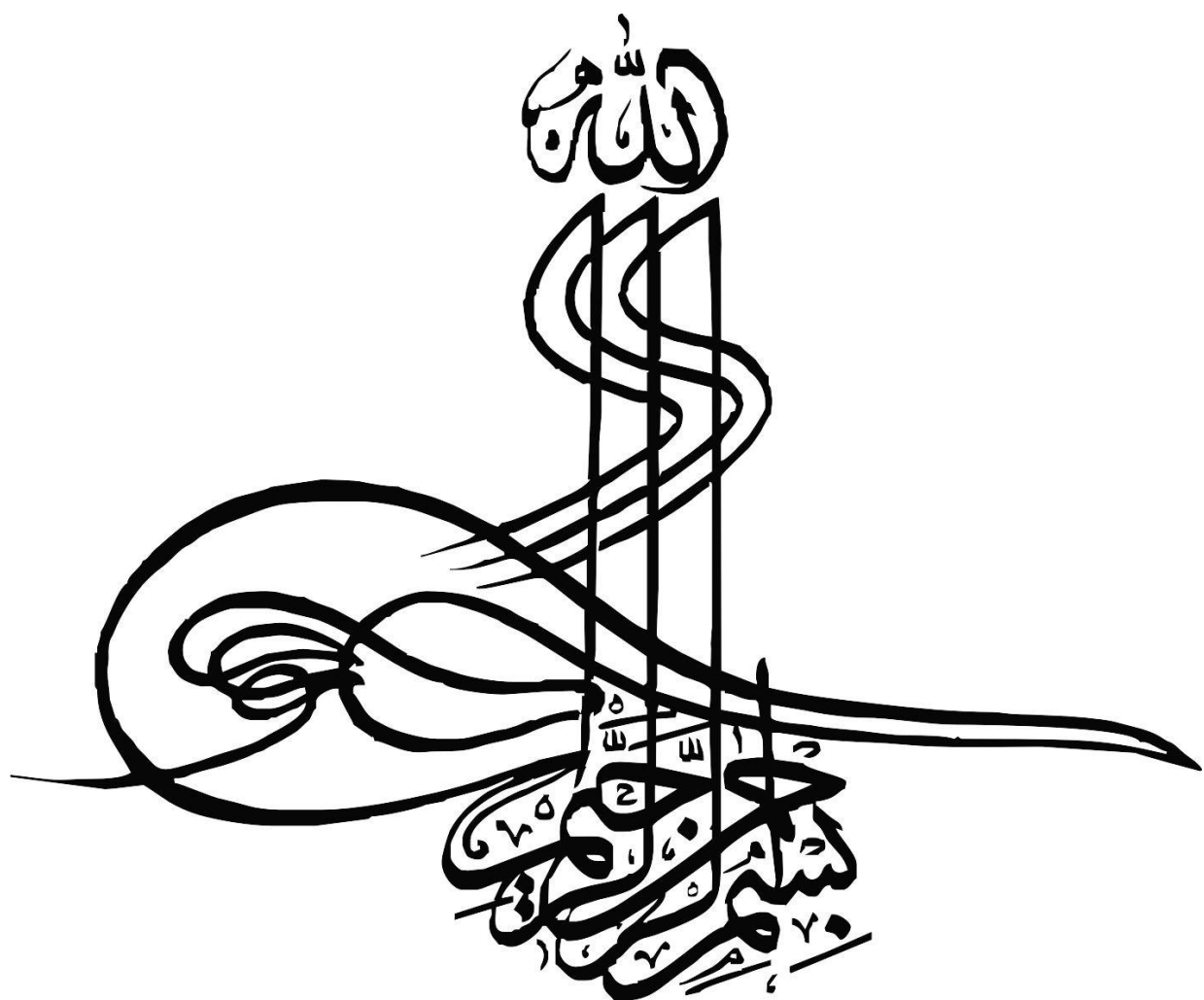






دانشگاه صنعتی شاهرود

دانشکده برق و رباتیک

پایان نامه کارشناسی ارشد

ردگیری هدف متحرک با استفاده از ماشین بردار پشتیبان

دانشجو:

رضا میغانی

استاد راهنما:

دکتر حیدر طوسی‌ان شاندیز

شهریور ۹۲

شماره : ۱۰۸۴/آ.ت.ب

تاریخ : ۹۲/۰۶/۲۳

ویرایش : ———

بسمه تعالی



مدیریت تحصیلات تکمیلی

فرم شماره (۶)

فرم صورتجلسه دفاع پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای :
رضا میغانی رشته : برق گرایش : کنترل
تحت عنوان : رهگیری هدف متحرک با استفاده از ماشین بردار پشتیبان
که در تاریخ ۹۲/۰۶/۲۳ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح زیر است :

☐ مردود	☐ دفاع مجدد	☒ قبول (با درجه : <u>خوب</u> امتیاز ۹۲۰)
--------------------------------	------------------------------------	--

۲- بسیار خوب (۱۸/۹۹ - ۱۸)

۱- عالی (۲۰ - ۱۹)

۴- قابل قبول (۱۵/۹۹ - ۱۴)

۳- خوب (۱۷/۹۹ - ۱۶)

۵- نمره کمتر از ۱۴ غیر قابل قبول

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنما	حسین طریقی	دانشیار	
۲- استاد مشاور	—	—	—
۳- نماینده شورای تحصیلات تکمیلی	سینا تریانی	استاد	
۴- استاد ممتحن	سید علی حسینی	دانشیار	
۵- استاد ممتحن	عزیزه الفی	استاد	

رئیس دانشکده :

تقدیر و تشکر

خدایا تو را سپاس

خدایا تو را سپاس به پاس توفیق آگاهی که به من ودیعه دادی. خدایا همچون گذشته به من قدرتی عطا کن تا به استقبال آینده‌ایی روشن پیش روم، صبر مرحمت فرما تا در سختی‌های روزگار دوباره بکوشم. امیدم، رویایی تازه است در انتظار، کمک کن تا بدان دست یابم.

با توجه به عنایات خاصه خداوند متعال، مساعدت و یاری خانواده عزیز، اساتید بزرگوار و دوستان مهربانم، اکنون که پایان نامه خود را به پایان رسانده‌ام بر خود لازم می‌دانم تا مراتب تشکر و قدردانی خود را ابراز نموده و کمال تشکر و امتنان را دارم.

در ابتدا **پدرم** که همواره بهترین پشتیبانم بوده است. **مادرم** که محکم‌ترین تکیه‌گاه زندگی‌ام است. سپس دوستان عزیزم که محبت‌های بی‌دریغشان هرگز فروکش نمی‌کند.

استاد ارجمند، آقای **دکتر حیدر طوسی‌ان شاندیز** که با راهنمایی‌های ارزنده، زحمات بی‌دریغ و تلاش‌های بی‌وقفه‌اشان در تمام مراحل انجام این پروژه، راهگشای اینجانب بودند.

همچنین از اعضای محترم هیئت داورى به خاطر حضور در جلسه دفاعیه و رفع اشکالات این پایان‌نامه تشکر و سپاسگزاری می‌کنم.

تعهد نامه

اینجانب رضا سیوانی دانشجوی دوره کارشناسی ارشد رشته برق کنترل دانشکده مهندسی برق و رباتیک دانشگاه صنعتی شاهرود نویسنده پایان نامه با عنوان :

بررسی صرف سرب با استفاده از ماشین بردار کشان

تحت راهنمایی آقای دکتر حیدر حبیبی متعهد می شوم :

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ: ۹۲/۶/۲۳

امضاء دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

* متن این صفحه نیز باید در ابتدای نسخه های تکثیر شده پایان نامه وجود داشته باشد .

چکیده:

یک الگوریتم ردیابی جسم پرنده براساس بکارگیری توأماً ویژگی‌های رنگ و بافت برای مدل کردن هدف و سپس بکارگیری الگوریتم جابجایی میانگین برای ردگیری هدف در این پایان‌نامه بکار گرفته شده است. علاوه بر ویژگی‌های مرسوم رنگ، از ویژگی‌های بافت یعنی ویژگی میانگین و الگوی دودویی محلی برای نمایش هدف بکار گرفته شده‌اند. همچنین از ماشین بردار پشتیبان برای جداسازی هدف از پس‌زمینه استفاده کرده‌ایم. برای آموزش ماشین بردار پشتیبان از ۴۰ تصویر از هدف مورد نظر که از زوایای متفاوت گرفته شده‌اند استفاده کرده‌ایم. در این تصاویر هدف به صورت دستی از پس‌زمینه جدا شده است. به منظور اینکه ممکن است برخی از پس‌زمینه‌های ممکن در طول ردیابی در تصاویر منتخب وجود نداشته باشند از گسترش پس‌زمینه استفاده می‌کنیم. چون داده‌های استخراج شده برای آموزش ماشین بردار پشتیبان زیاد است داده‌هایی که در تعیین فوق صفحه جداکننده نقش خاصی ندارند را حذف می‌کنیم و ماشین بردار پشتیبان را آموزش می‌دهیم. بعد از دریافت اولین فریم و آزمایش آن توسط ماشین بردار پشتیبان، آن را توسط یک چهارچوب نمایش می‌دهیم و از الگوریتم جابجایی میانگین برای ردگیری آن در فریم‌های بعد کمک می‌گیریم. نتایج بدست آمده کارایی الگوریتم پیشنهادی را در دقت و سرعت تأیید می‌کند.

کلمات کلیدی: ردگیری هدف، ماشین بردار پشتیبان، بینایی ماشین، الگوی دودویی محلی، جابجایی میانگین، گسترش پس‌زمینه

فهرست

فصل اول: مقدمه‌ای بر تشخیص و ردیابی اشیاء متحرک	۱
مقدمه	۲
۱-۱- اهداف پایان‌نامه	۴
۲-۱- ساختار پایان‌نامه	۴
فصل دوم: پیشینه تحقیق	۵
مقدمه	۶
۱-۲- طبقه‌بندی روش‌های ردیابی	۶
۱-۱-۲- ردیابی نقطه‌ای	۸
۲-۱-۲- ردیابی بر اساس هسته	۹
۳-۱-۲- ردیابی سیاه‌نما	۱۹
۴-۱-۲- روش‌های تفریق فریم و تفریق زمینه	۲۲
۲-۲- حذف نویز	۲۸
۳-۲- ویژگی‌های بکار رفته در ادبیات تحقیق	۳۲
۳-۲-۱- رنگ	۳۲
۳-۲-۲- بافت	۳۲

۳۳	 ۲-۳-۳- لبه
۳۳	 ۲-۴- پیچیدگی ها و موانع مساله ردیابی
۳۴	 ۲-۴-۱- کاهش ابعاد
۳۴	 ۲-۴-۲- حرکت دوربین
۳۴	 ۲-۴-۳- تغییرات روشنایی صحنه
۳۴	 ۲-۴-۴- ساختار غیر صلب شی
۳۵	 ۲-۴-۵- حرکات پیچیده و ناگهانی شی
۳۵	 ۲-۴-۶- تغییر شکل و اندازه شی
۳۵	 ۲-۴-۷- پوشیدگی شی
۳۶	 ۲-۴-۸- التزامات پردازش به هنگام
۳۶	 ۲-۵- جمع بندی
۳۷	 فصل سوم: ماشین بردار پشتیبان
۳۸	 مقدمه
۳۸	 ۳-۱- ابر صفحه جداساز
۳۹	 ۳-۲- صفحه بهینه برای الگوهای خطی تفکیک پذیر
۴۲	 ۳-۳- صفحه بهینه برای الگوهای خطی تفکیک ناپذیر

۴۴	۴-۳- ماشین بردار پشتیبان برای داده‌های غیر خطی
۴۵	۴-۳-۱- ضرب داخلی کرنل
۴۷	۴-۳-۲- تئوری مرسر
۴۸	۴-۳-۳- طراحی بهینه ماشین بردار پشتیبان
۴۹	۴-۳-۵- نقاط قوت و ضعف ماشین بردار پشتیبان
۵۰	۴-۳-۶- یافتن مقادیر بهینه برای سیگما و C
۵۱	۴-۳-۷- ماشین بردار پشتیبان برای داده‌های خیلی بزرگ
۵۵	فصل چهارم: روش پیشنهادی و جزئیات پیاده سازی
۵۶	مقدمه
۵۶	۴-۱- آموزش ماشین بردار پشتیبان
۵۷	۴-۲- ردگیری هدف
۵۹	۴-۳- ویژگی‌های بکار رفته در پایان نامه
۵۹	۴-۳-۱- ویژگی رنگ
۵۹	۴-۳-۲- ویژگی بافت
۶۲	۴-۴- جابجایی میانگین
۶۳	۴-۵- روش خوشه‌بندی k_mean

۶-۵- گسترش پس زمینه	۶۶
۸-۴- بروز رسانی تغییرات در اندازه و مدل شی	۶۸
فصل پنجم: نتایج شبیه سازی، نتیجه گیری و پیشنهادات	۷۱
مقدمه	۷۲
۱-۵- بدست آوردن مسیر واقعی	۷۲
۲-۵- استخراج ویژگی	۷۴
۳-۵- ردگیری هدف	۷۴
۴-۵- نتیجه گیری	۷۸
۵-۵- پیشنهادات	۷۸
منابع	۷۹

فهرست شکل‌ها

- شکل (۱-۲) روش‌های مختلف مدل کردن شیء. به ترتیب از سمت چپ به راست، مدل کردن شیء توسط نقاط، مدل کردن توسط هسته، مدل کردن توسط سیاه نما و استفاده از روش تفریق پس‌زمینه برای مدل کردن ۸
- شکل (۲-۲) طبقه‌بندی روش‌های ردیابی ۸
- شکل (۳-۲) نقاط آشکارسازی توسط آشکارسازهای نقاط: روش‌های (a) Harris, (b) KLT, (c) SIFT ۹
- شکل (۴-۲) نمایش مشکل لغزش در بروز رسانی قالب ۱۱
- شکل (۵-۲) نتایج حاصل از الگوریتم ردیابی جابجایی میانگین ۱۲
- شکل (۶-۲) شکل بالا تصویر مادون قرمز از یک کامیون را نشان می‌دهد. مقادیر ویژگی‌های شیء و زمینه اطراف آن برای محاسبه نسبت لگاریتم احتمال هیستوگرام‌های شیء و زمینه استفاده شد که پیکسل‌های شیء را به مقادیر مثبت و پیکسل‌های زمینه را به مقادیر منفی نگاشت می‌دهد. به این ترتیب شیء را از زمینه جدا می‌کند ۱۴
- شکل (۷-۲) نتایج ردیابی در [۱۰]، تعدادی فریم از یک فیلم به همراه نقشه اطمینان قسمت مربوط به قسمت انتخاب شده توسط مستطیل ۱۵
- شکل (۸-۲) هنگامی که زمینه رنگی نزدیک به هدف داشته باشد الگوریتم ردگیری دچار اشتباه می‌شود. در شکل بالا هدف بعد از رسیدن به نقاط سفید در زمینه اشتباه این نقاط در نظر گرفته می‌شود ۱۶
- شکل (۹-۲) سه فریم از دنباله ویدیوی هستند و d امتیاز ماشین بردار پشتیبان به شیء مورد ردیابی در طول ردیابی ۱۸

- شکل (۲-۱۰)- تعدادی فریم از دنباله ویدئویی، به همراه مدل هدف مورد ردیابی ۲۰
- شکل (۲-۱۱)- نتایج الگوریتم ردیابی کانتور ۲۱
- شکل (۲-۱۲) خروجی الگوریتم تفاضل فریم های متوالی ۲۴
- شکل (۲-۱۳) پردازش حذف پس زمینه الف) تصویر رنگی جاری و پس زمینه. ب) خروجی پردازش حذف پس زمینه ۲۷
- شکل (۲-۱۴) پردازش حذف تشخیص اشتباه. الف) تصویر جاری ب) پردازش پس زمینه بدون حذف تشخیص اشتباه. ج) پردازش حذف اشتباه با در نظر گرفتن یک نقطه در همسایگی د) پردازش حذف اشتباه با در گرفتن همه مولفه متصله همسایگی ۳۱
- شکل (۳-۱): صفحه جداساز بهینه با حداکثر مقدار حاشیه ۳۹
- شکل (۳-۲) : صفحه جداساز و حاشیه ها ۴۰
- شکل (۳-۳): سیستم های خطی جداناپذیر با میزان خطای ϵ_i ۴۲
- شکل (۳-۴): نگاشت غیر خطی از فضای ورودی به فضای ویژگی ۴۵
- شکل (۳-۵)- نتایج حاصل برای انتخاب بهینه پارامترهای ماشین بردار پشتیبان ۵۰
- شکل (۳-۶)- آشکارسازی لبه در مسئله کلاس بندی ۵۲
- شکل (۳-۷)- کاهش داده های آموزشی ۵۳
- شکل (۳-۸)- داده های آموزشی در ماشین بردار پشتیبان استاندارد و روش ارائه شده ۵۴
- شکل (۳-۹)- مقایسه درصد درستی ماشین بردار پشتیبان استاندارد و روش ارائه شده ۵۴
- شکل (۳-۱۰)- مقایسه زمان لازم جهت آموزش ماشین بردار پشتیبان استاندارد و روش ارائه شده ۵۴

- شکل (۴-۱) - مراحل مربوط به آموزش ماشین بردار پشتیبان ۵۷
- شکل (۴-۲) - فلوچارت مربوط به مرحله تشخیص و ردگیری هدف ۵۸
- شکل (۴-۳) - یک ماسک 3×3 ۶۰
- شکل (۴-۴) - ماسک استفاده شده در پایان نامه ۶۱
- شکل (۴-۵) - یک مثال از $LBP_{8,1}$ ۶۲
- شکل (۴-۶) - روند به صفر رسیدن تعداد داده‌های اشتباه خوشه‌بندی شده ۶۵
- شکل (۴-۷) - خوشه‌بندی داده‌ها در پنج کلاس ۶۵
- شکل (۴-۸) - آزمایش تعدادی داده در فضای ویژگی قبل از گسترش پس‌زمینه ۶۶
- شکل (۴-۹) - داده‌های استخراج شده از عکس‌های انتخابی و داده‌های بدست آمده از گسترش پس‌زمینه ۶۸
- شکل (۴-۱۰) - حالاتی که برای به روز رسانی اندازه پیش می‌آید. ۶۹
- شکل (۵-۱) - تصویر به همراه شکل باینری آن و مرکز ثقل آن ۷۳
- شکل (۵-۲) - سطر اول تصاویر منتخب برای آموزش. سطر دوم تصاویر متناظر با سطر اول که هدف بصورت دستی جدا شده. تصاویر باینری مربوط که دارای برچسب یک و صفر است ۷۳
- شکل (۵-۳) - ویژگی‌های استخراج شده از تصاویر شکل (۵-۲)، ویژگی‌های رنگ R، G، B، الگوی دودویی محلی و میانگین به ترتیب در سطرهای یک تا پنج ۷۴
- شکل (۵-۴) - به ترتیب از چپ به راست، اولین فریم از فیلم، تصویر باینری معادل آن که خروجی ماشین بردار پشتیبان است و قسمتی از تصویر که به عنوان هدف تشخیص داده شد. ۷۵

شکل (۵-۵). چهارچوب کم‌رنگ مربوط به فریم قبل و چهارچوب تیره مربوط به فریم فعلی ۷۵

شکل (۶-۵). فریم‌های ۲۰، ۵۰، ۱۰۰ و ۱۵۰ به همراه چهارچوب باینری مربوط به آن ۷۶

شکل (۷-۵) اختلاف مسیر طی شده‌ی هدف در الگوریتم‌های متفاوت با مسیر اصلی ۷۷

فهرست جداول

جدول (۱-۳) خلاصه از ضرب داخلی کرنل‌ها ۴۹

جدول (۱-۵). مقایسه زمان اجرای الگوریتم پیشنهادی با الگوریتم‌های معرفی شده در بالا ۷۷

فصل اول

مقدمه ای بر تشخیص و ردیابی اشیاء متحرک

مقدمه

با گسترش فن‌آوری، نیاز به سیستم‌های هوشمند و مستقل از کاربر با دقت، سرعت و قابلیت بیشتر و بالاتر روز به روز بیشتر می‌شود. از جمله این نیازها، نیاز به یک ماشین جهت تشخیص و ردیابی است، که به آن ماشین بینایی می‌گویند. در بسیاری از کاربردهای بینایی ماشین، از خروجی سیستم تشخیص و ردیابی به عنوان ورودی در پردازش‌های مرتبه بالاتر مانند تعبیر و تفسیر حرکت، شمارش تعداد اشیاء، تشخیص نوع رفتار و نظایر آن استفاده می‌شود.

در ردیابی اشیاء متحرک، قبل از ردگیری لازم است شیء متحرک از پس زمینه جدا شود. باید توجه داشت، که تشخیص اشیاء متحرک از یک پس‌زمینه، ردیابی نمی‌باشد اما در ردیابی اشیاء مورد استفاده قرار می‌گیرد.

بسته به کاربرد سیستم ردیاب، متغیرهایی چون محل و موقعیت، سرعت، اندازه و جهت، نحوه قرارگرفتن^۱، شکل و سایر داده‌های مربوط به یک شیء می‌توانند به عنوان خروجی‌های سیستم ردیاب تلقی شوند. بنابراین بطور کلی، می‌توان فرآیند ردیابی را عمل بدست آوردن، استنتاج و تخمین تغییرات زمانی و مکانی شیء یا حالت‌های شیء یا هدف در طول دنباله ویدئویی بر اساس اندازه‌گیری‌ها و مشاهدات تعریف کرد.

از لحاظ کاربردی، روش‌های ردیابی را به دو دسته می‌توان تقسیم نمود.

(۱) کاربردهای مراقبتی، که در آن پردازش بلادرنگ^۲ و خودکار^۳ نیاز اصلی می‌باشد و حتی

مقداری عدم دقت در نمایش شیء نیز قابل قبول می‌باشد.

^۱- Pose

^۲- Real-time

^۳- Automatic

۲) کاربرد فشرده‌سازی ویدیو، که در آن دقت در نمایش شی و استخراج نقاط مربوط به شی، بسیار مهم می‌باشد. به گونه‌ای که حتی ممکن است این میزان از دقت به عدم پردازش بلادرنگ و حتی دخالت کاربر در استخراج شی یا اعمال شرایط کنترل‌شده خاص منجر شود.

با وجودیکه تشخیص و ردیابی اشیاء چندین سال است که مورد مطالعه و تحقیق قرار گرفته است، اما هنوز یک روش دقیق، مقاوم با کارایی بالا در بسیاری از کاربردها وجود ندارد. میزان سختی مسئله به مقدار زیادی به چگونگی تعریف ما از شی مورد ردیابی برمی‌گردد. اگر فقط شمار کمی از ویژگی‌های تصویر مانند رنگ شی برای نمایش شی مورد استفاده قرار گیرد، تشخیص شی با تعیین نقاطی که دارای رنگ یکسان هستند، امری نسبتاً ساده خواهد بود. از طرف دیگر تشخیص، شناسایی و ردیابی چهره یک فرد مشخص که دارای جزئیات زیاد، حالت‌ها و شرایط نوری متفاوت می‌باشد، بسیار سخت خواهد بود.

تغییرات در تصاویر گرفته شده از چهار منبع حاصل می‌شود :

۱- تغییرات در حالت^۱ شی

۲- تغییر شکل^۲ شی

۳- تغییرات در شرایط نوری محیط^۳

۴- انسداد یا همپوشانی نسبی (جزیی) یا کلی

¹- Pose

²- Deformation

³- Illumination

۱-۱- اهداف پایان نامه:

در این پایان نامه هدف، ردیابی یک نوع خاص جسم پرنده براساس ویژگی‌های رنگ و بافت و بکارگیری الگوریتم جابجایی میانگین برای ردیابی است. این روش قابلیت ردیابی شیء با وجود متحرک بودن دوربین را دارد. و در برابر همپوشانی جزئی نیز مقاوم است. از محدودیت‌های روش پیشنهادی این است که باید در هر فریم قسمتی از هدف داخل مستطیلی که به عنوان مدل شیء در نظر گرفته شده باشد. برای غلبه بر این محدودیت می‌توان از روش‌های پیش‌بینی برای یافتن مسیر بعدی هدف استفاده کرد و در آن نقاط دنبال هدف باشیم.

۱-۲- ساختار پایان نامه:

این پایان نامه دارای پنج فصل است. در فصل دوم پیشینه تحقیق مرور خواهد شد. در فصل سوم الگوریتم‌های بکار رفته در پایان نامه اعم از روش‌های استخراج ویژگی، روش جابجایی میانگین و روش برورسانی ابعاد مستطیل مورد بررسی قرار خواهد گرفت. در فصل چهارم به بررسی ماشین بردار پشتیبان و چگونگی تنظیم پارامترهای آن و کاهش ویژگی‌های بکار رفته می‌پردازیم. در فصل پنجم نیز نتایج تحقیق مورد بررسی قرار خواهد گرفت. در انتهای فصل پنجم نتیجه گیری و پیشنهادات نیز آورده شده است.

فصل دوم

پیشینه تحقیق

مقدمه:

در این فصل پیشینه‌ی تحقیق روش‌هایی که در حوزه ردیابی اشیاء در حوزه‌ی دو بعدی ارائه شده‌اند بررسی می‌شوند و بالاخص روش‌های مدل‌سازی و ردیابی شیء که مرتبط با کار انجام شده است با تمرکز بیشتری مورد بررسی قرار می‌گیرند. در ادامه، ویژگی‌های مختلفی که در ردیابی شیء بیشتر مورد استفاده قرار می‌گیرند بررسی می‌شوند و در انتها پیچیدگی‌ها و موانع اصلی برای مساله ردیابی اشیاء بیان خواهد شد.

۲-۱- طبقه‌بندی روش‌های ردیابی

ادبیات ردیابی بعلاوه رویکردهای بسیار زیادی که به کاربردهای مختلف وجود دارد بسیار وسیع است. تا کنون متخصصان زیادی از گرایش‌های الکترونیک، مخابرات، کامپیوتر، کنترل با رویکردهای مختلفی را برای حل مسئله‌ی ردیابی شیء ارائه داده‌اند. به علت گستردگی زیاد رویکردهای ارائه شده در خصوص ردیابی شیء تا کنون تحقیقات جامع زیادی بر ادبیات ردیابی شیء صورت نگرفته است. جامع‌ترین تحقیق انجام شده در حوزه‌ی ردیابی دو بعدی در [1] صورت گرفته است. تحقیق دیگری اخیراً در [2] صورت گرفته که بیشتر مبنی بر پژوهش قبلی است. در این بخش روش‌هایی که در حوزه‌ی ردیابی کلی اشیاء ارائه شده در نظر گرفته شده است و روش‌هایی که تنها برای یک شیء خاص بکار می‌روند مورد تحقیق قرار نگرفته‌اند. برای مثال آن دسته از روش‌هایی که از جنبش‌شناسی بشر^۱ به عنوان اساس پیاده‌سازی استفاده می‌کنند و یا روش‌هایی که تنها برای اشیاء مفصلی بکار رفته‌اند و قابل استفاده برای اشیاء غیر مفصلی نیستند مد نظر قرار نمی‌گیرند. همچنین در این تحقیق تنها روش‌های مطرح شده که در حوزه‌ی ردیابی دو بعدی مورد استفاده قرار گرفته‌اند، مورد بررسی قرار می‌گیرد.

¹ human kinematics

هدف این بخش دسته بندی الگوریتم‌های شناخته شده تر و بیان نقاط ضعف و قدرت آن‌ها ، با تمرکز بر روش‌های بکار رفته در پایان‌نامه است. ابتدا روش‌های متداول برای مدل کردن و ردیابی شیء بیان می‌شود. در ادامه از آنجا روش بکار گرفته شده، ردیابی بر اساس ویژگی‌های استخراج شده از پیکسل‌های تصویر است بطور مختصر ویژگی‌های مختلفی که در ادبیات ردیابی شیء بیشتر بکار رفته‌اند بررسی خواهیم کرد. لازم به ذکر است که ویژگی بافت و رنگ که در این پایان‌نامه از آن استفاده شده است با جزئیات بیشتری در فصل بعد مورد بررسی قرار می‌گیرد.

بطور خلاصه بازنمایی‌های متداول بکار رفته برای مدل کردن و ردیابی شیء عبارتند از [1]

-ردیابی نقاط^۱: در این نوع ردیابی، شیء می‌تواند توسط نقاط خاصی از آن که دارای ویژگی‌های متمایزی است مدل گردد و این نقاط به عنوان مدل شیء در فریم‌های متوالی ردیابی می‌گردد.

-ردیابی بر اساس هسته^۲: شیء می‌تواند توسط یک شکل اولیه مانند بیضی یا مستطیل که شیء در آن قرار دارد متمایز گردد و یک قالب یا یک هسته تشکیل شده از ویژگی‌های شیء به عنوان مدل شیء در نظر گرفته شود.

-ردیابی سیاه‌نما^۳: شیء می‌تواند توسط محدوده‌ی دقیق آن مدل گردد. بازنمایی کانتور، مرزهای یک شیء را تعیین و مشخص می‌کند و ناحیه درون کانتور، سیاه نمای شیء نامیده می‌شود. در این دسته از رویکرد هدف ردیابی سیاه‌نمای شیء در فریم‌های متوالی است.

-روش‌های تفریق فریم^۴ یا زمینه^۱: یافتن ناحیه‌شی و ردیابی آن با ساختن مدلی از زمینه و تفاضل فریم‌ها از آن انجام می‌گیرد یا مستقیماً با تفاضل فریم‌ها شیء استخراج خواهد شد. در شکل (۲-۱)

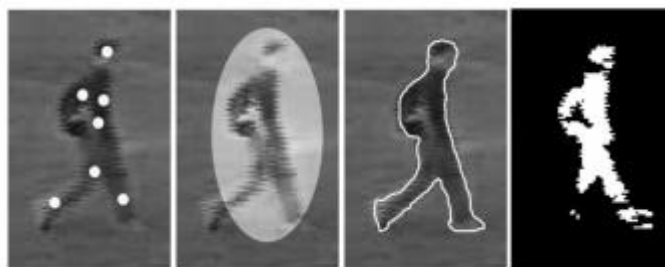
¹ Point tracking

² Kernel tracking

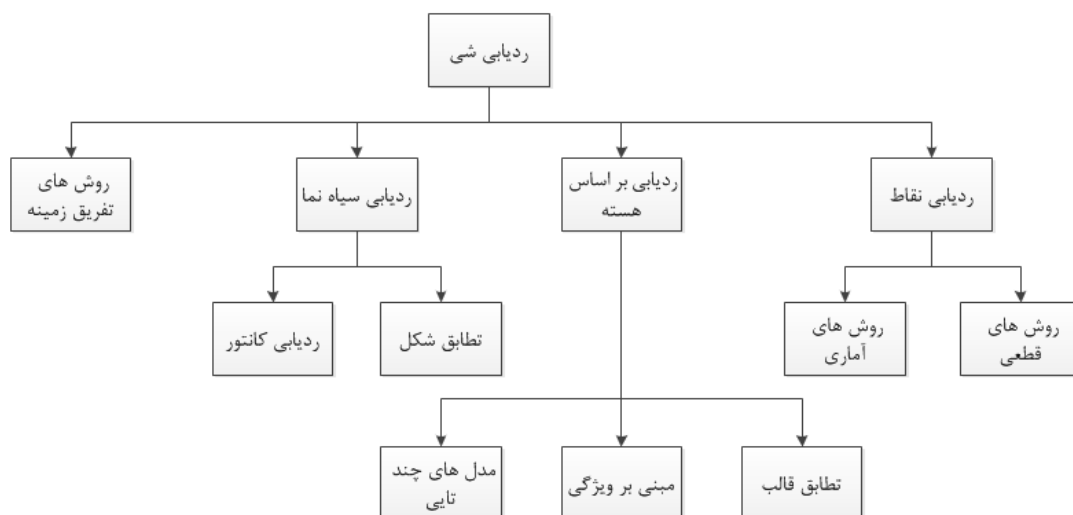
³ Silhouette tracking

⁴ Frame differencing

روش‌های متداول مدل کردن شیء در کاربردهای ردیابی نمایش داده شده است. در شکل (۲-۲) یک طبقه‌بندی از روش‌های ردیابی متداول که توضیح آن‌ها در ادامه می‌آید نمایش داده شده است. روش بکار گرفته شده در پایان‌نامه جز روش‌های ردیابی بر اساس هسته است. بنابراین با تفضیل بیشتری به مطالعه رویکردهای این دسته خواهیم پرداخت.



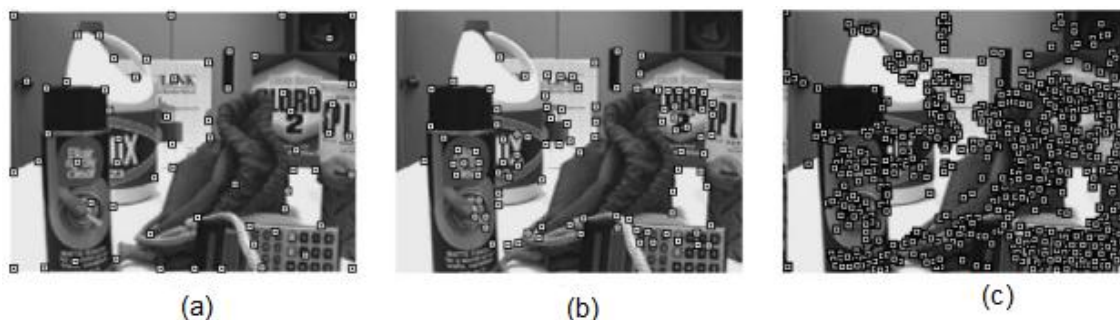
شکل (۲-۱) روش‌های مختلف مدل کردن شیء. به ترتیب از سمت چپ به راست، مدل کردن شیء توسط نقاط، مدل کردن توسط هسته، مدل کردن توسط سیاه نما و استفاده از روش تفریق پس‌زمینه برای مدل کردن [۱]



شکل (۲-۲) - طبقه‌بندی روش‌های ردیابی برگرفته شده از [۱]

۲-۱-۱-ردیابی نقاط

ردیابی می‌تواند بصورت تطبیق اشیاء بازنمایی شده با نقاط در طول فریم‌ها فرموله شود. این نقاط ابتدا توسط آشکارسازهای نقاط مشخص می‌گردد. آشکارسازهای نقاط برای یافتن نقاطی در تصویر استفاده می‌شوند که دارای بافت‌های پرمعنا در مکان‌هایشان باشند. یک نقطه‌ی مطلوب آن است که نسبت به تغییرات روشنایی و زاویه‌ی دید دوربین تغییرناپذیر باقی بماند. این نقاط مدت نسبتاً طولانی است که در کاربردهای حرکت و ردیابی استفاده می‌شوند. آشکارسازهای نقاط بکار رفته شامل آشکارسازهای Moravec, Harris, KLT, SIFT^۱، است. یک ارزیابی مقایسه‌ای از آشکارسازی نقاط در [3] ارائه شده است. بطور کلی آشکارساز نقاط SIFT از اکثر آشکارسازهای نقاط نتایج بهتر و مفیدتری دارد و در برابر تغییر شکل‌های تصویر انعطاف پذیری بیشتری دارد. مثالی از عملکرد سه آشکارساز نقاط SIFT, Harris, KLT در شکل (۳-۲) نشان داده شده است. پس از آشکارسازی نقاط، این نقاط باید در فریم‌های متوالی به هم تطبیق داده شوند. تطبیق نقاط در فریم‌های متوالی مخصوصاً در حضور پوشیدگی مساله‌ی پیچیده‌ای است.



شکل (۳-۲) - نقاط آشکارسازی توسط آشکارسازهای نقاط: روش‌های $SIFT(c)$, $KLT(b)$, $Harris(a)$ [۱]

^۱Scale Invariant Feature Transform

۲-۱-۲) ردیابی بر اساس هسته

شیء می‌تواند توسط یک هسته اولیه مانند بیضی یا یک مستطیل که شیء در آن قرار گرفته است مدل کرد و یک قالب یا یک هسته متشکل از ویژگی‌های شیء بعنوان مدل شیء در نظر گرفت. ردیابی بر اساس هسته معمولاً با محاسبه حرکت شیء که توسط یک شکل اولیه بازنمایی شده است از فریمی به فریم بعد انجام می‌گیرد. الگوریتم‌های بر اساس هسته را می‌توان به سه دسته‌ی بر اساس قالب، بر اساس ویژگی و بر اساس مدل‌های چندتایی تقسیم کرد.

۲-۱-۲-۱) روش‌های تطابق قالب:

تطابق قالب‌ها یکی از متداول‌ترین و قدیمی‌ترین رویکردهای ردیابی شیء است که به علت سادگی نسبی بطور گسترده‌ای مورد استفاده قرار گرفته است و تاریخ آن به الگوریتم [4] Lucas-kanade در سال ۱۹۸۱ بر می‌گردد. قالب‌ها توسط شکل‌های هندسی ساده ساخته می‌شوند. در این روش ابتدا توسط یک قالب (که می‌تواند یک شکل هندسی اولیه باشد مانند مستطیل یا بیضی) که اغلب شیء درون آن قرار گرفته است مدل می‌گردد. تطابق قالب یک رویکرد جستجوی سراسری در تصویر (I_w) برای یافتن یک ناحیه مشابه با قالب شیء (Q_t) است. مکان قالب در تصویر فعلی با معیار شباهت محاسبه می‌شود. الگوریتم‌های تطابق قالب ساده معمولاً از یک قالب ثابت در تمامی دنباله‌ی تصویر استفاده می‌کنند که معمولاً از شدت تصویر یا ویژگی‌های رنگ برای ساخت قالب استفاده می‌شود. این استراتژی تا زمانی می‌تواند معتبر باشد که ظاهر شیء تغییر نکند. یک راه حل برای این مسئله بروز رسانی قالب در طول ردیابی است به این معنی که بهترین ناحیه‌ی تطبیق یافته در فریم فعلی به عنوان قالب جدید برای فریم بعدی در نظر گرفته شود. مسئله‌ی مهمی که در اینجا وجود دارد این است که چطور قالب را در طول زمان بروز رسانی کنیم تا مدل خوبی از شیء مورد ردیابی باقی بماند. اگر قالب بطور مناسبی در طول بروز رسانی نگردد ممکن است جابجایی کوچکی نسبت به مکان اصلی شیء در فریم بعد بوجود بیاید و این خطاها می‌تواند در طول زمان جمع شده و باعث انحراف از شیء

مورد نظر برای ردیابی گردد که مساله لغزش، مساله شده‌ای در رویکردهای بر اساس بروز رسانی قالب‌هاست [۵]. این مساله در شکل (۲-۴) به خوبی نشان داده شده است.



شکل (۲-۴) نمایش مشکل لغزش در بروز رسانی قالب [۵]

یک محدودیت روش تطابق قالب‌ها هزینه‌ی بالای محاسباتی آن به علت رویکرد جستجوی سراسری است. برای کاهش هزینه‌ی محاسباتی محققان معمولاً جستجوی شیء را به همسایگی مکان شیء محدود می‌کنند. یک مزیت قالب این است که علاوه بر اطلاعات ظاهری شیء، اطلاعات مکانی شیء را نیز در خود دارد. از سوی دیگر یکی از مشکلات استفاده از قالب شیء این است که اطلاعات ظاهر شیء را تنها از یک زاویه دید در خود دارد بنابراین استفاده از قالب تنها هنگامی مناسب است که ظاهر شیء در طول ردیابی تغییرات قابل ملاحظه‌ای نداشته باشد. یکی دیگر از مشکلات استفاده از قالب که در اغلب روش‌های ردیابی بر اساس هسته نیز وجود دارد این است که به علت اینکه مدل شیء بطور دقیقی توسط شکل‌های هندسی ابتدایی مدل نمی‌گردد اغلب قسمتی از اطلاعات زمینه نیز به همراه شیء مدل می‌شود و این باعث ضعیف شدن مدل شیء می‌گردد.

۲-۲-۱-۲- روش‌های مبنی بر ویژگی

بجای قالب، بازنمایی‌های دیگری می‌تواند در ردیابی استفاده کرد. برای مثال هیستوگرام رنگ می‌تواند با استفاده از پیکسل‌های درون ناحیه‌ی مستطیلی یا بیضوی محاسبه گردد و به عنوان مدل شیء در نظر گرفته شود. در [۶] نویسندگان مدل شیء را توسط میانگین رنگ و پیکسل‌های درون ناحیه‌ی درون مستطیلی ساختند و برای کاهش پیچیدگی محاسباتی در فریم بعد شیء را در ۸ همسایگی اطراف شیء

جستجو کردند. شباهت بین مدل شی (M) و مکان‌های فرضی (H) با ارزیابی میانگین رنگ‌های (M,H) محاسبه می‌گردد و بهترین آن‌ها به عنوان مکان فعلی شی انتخاب می‌گردد. در [۷] نویسندگان از هیستوگرام وزن داده شده‌ی محاسبه شده از یک ناحیه بیضوی برای بازنمایی شی استفاده کرده‌اند و بجای انجام جستجوی سراسری آن‌ها از پروسه‌ی جابجایی میانگین برای مکان یابی شی استفاده کردند. روش جابجایی میانگین با مقایسه‌ی هیستوگرام شی Q و پنجره‌ی اطراف مکان شی P با چندین بار تکرار (معمولاً ۵ تا ۶ تکرار) شباهت آن‌ها را بیشینه می‌کند. معیار شباهت هیستوگرام‌ها توسط ضرایب باهاتاچاریا طبق رابطه (۲-۲) تعریف می‌گردد.

$$(۱-۲) \quad \text{ضریب باهاتاچاریا} = \sum_{u=1}^b P(u)Q(u)$$

که b تعداد بین‌های هیستوگرام (تعداد نقاط طول نمودار هیستوگرام مثلاً برای تصاویر خاکستری تعداد بین‌ها ۲۵۶ و برای تصاویر باینری دو است) رنگ است. برای ساخت هیستوگرام، نویسندگان از یک وزندهی تعریف شده با هسته فاصله‌ای استفاده کردند که به پیکسل‌های نزدیکتر به مرکز شی وزن بیشتری می‌دهد.



شکل (۲-۵)- نتایج حاصل از الگوریتم ردیابی جابجایی میانگین [۷]

یکی از نویسندگان مقاله‌ی [۷]، در مقاله‌ای دیگر [۸] از هیستوگرام مشترک فاصله-رنگ (بجای هیستوگرام رنگ) با رویکرد جابجایی میانگین استفاده کرد. مثالی از روش ردیابی جابجایی میانگین ارائه شده در مقاله [۷] در شکل (۲-۵) نمایش داده شده است. یک مزیت واضح ردیابی جابجایی میانگین نسبت به رویکرد تطابق قالب استاندارد حذف جستجوی سراسری کل تصویر است. با این وجود ردیاب جابجایی میانگین نیاز دارد که در فریم بعدی حداقل قسمتی از شی درون ناحیه بیضوی اولیه قابل مشاهده باشد. برای حذف چنین نیازی یک راه حل امکان‌پذیر استفاده از فیلتر کالمن یا فیلتر ذره‌ای برای پیش‌بینی مکان شی در فریم بعدی است. با داشتن مکان شی و تعریف کردن سرعت و شتاب حرکت مرکز ثقل، این فیلترها مرکز ثقل شی را طوری تخمین می‌زنند که در فریم بعد احتمال مشاهده‌ی شی درون هسته افزایش می‌یابد. [۷]

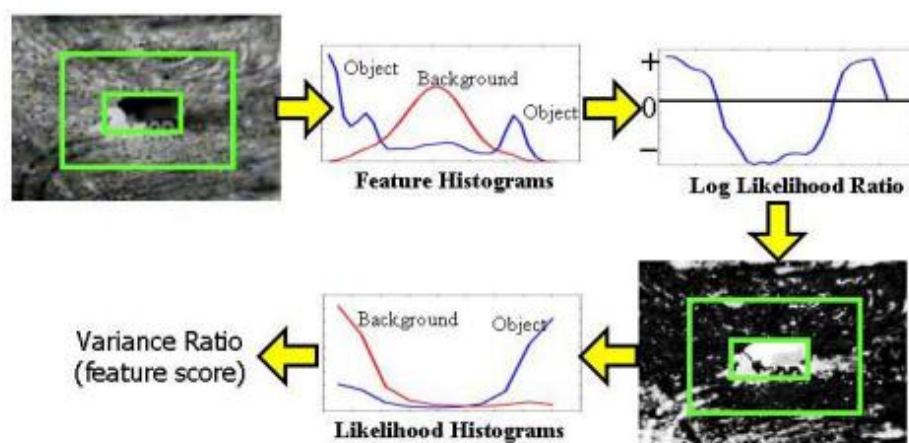
در روش‌های مرور شده تا اینجا همانند دیگر رویکردهای ردیابی بر اساس هسته، مدل شی نمی‌تواند بطور دقیقی توسط شکل‌های هندسی مدل می‌گردد. در [۹-۱۰] نویسندگان با تعریف جدی‌تر از ردیابی شی سعی در حل این مشکل نمودند. به این صورت که نویسندگان مساله ردیابی شی را بصورت جداسازی شی از زمینه‌ی نزدیک به شی را بصورت جداسازی شی از زمینه‌ی نزدیک به شی (که اطراف شی قرار دارد و نه کل زمینه) در نظر گرفتند. آنها با استفاده از رابطه (۲-۲) که نسبت لگاریتم احتمال هیستوگرام‌های ویژگی‌های ناحیه‌ی شی $H_o(i)$ و زمینه اطراف شی $H_b(i)$ را محاسبه می‌کند سعی در جدایی شی مورد ردیابی از زمینه اطرافش در هر فریم کردند و سپس با رویکرد جابجایی میانگین شی را ردیابی کردند.

$$L = \log \frac{\max(H_o(i))}{\max(H_b(i))} \quad (2-2)$$

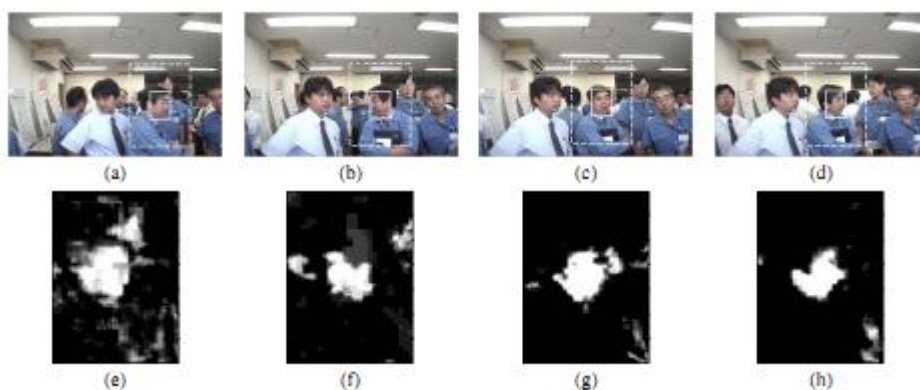
که اندیس i بیانگر تعداد المان‌های هیستوگرام ویژگی‌ها است. توصیفی از این رویکرد در شکل (۲-۶) نمایش داده شده است. اما ضعف این روش این است که اولاً تشکیل هیستوگرام‌ها دارای حجم

محاسباتی بالایی بوده و با افزایش ویژگی‌ها دچار بزرگی ابعاد شده و عملاً بسیار زمانبر می‌گردد، ثانیاً اگر شی دیگری با ویژگی‌های مشابه با هدف قرار داشته باشد مدل هدف تضعیف می‌گردد و دیگر اینکه همانند ردیاب جابجایی میانگین نیاز دارد که در فریم بعدی حداقل قسمتی از شی درون ناحیه انتخاب شده‌ی اولیه قابل مشاهده باشد.

نویسنده مقاله [۱۰] ردیابی شی را بعنوان یک مسئله‌ی کلاس‌بندی دو کلاسه در نظر گرفت و با ساخت یک تابع جداساز متغیر با زمان، توزیع‌های ویژگی شی و زمینه را در هر فریم از هم جدا کرد. او با استفاده از آموزش ویژگی‌های ناحیه‌ی شی و زمینه اطراف شی به یک مجموعه از کلاس‌بندی‌های ضعیف و سپس ترکیب آنها نقشه‌ی احتمال شی را تشکیل داد و برای یافتن پیک نقشه‌ی احتمال و بنابراین مکان جدید شی از رویکرد جابجایی میانگین شی استفاده کرد. ضعف این روش این است که از آستانه‌گیری‌های متعددی استفاده کرده است و باید از قبل دانسته شود تا برای تعیین آستانه‌ها مورد استفاده قرار گیرد. مثالی از این رویکرد ردیابی در شکل (۲-۷) نمایش داده شده است.



شکل (۲-۶)- شکل بالا تصویر مادون قرمز از یک کامیون را نشان می‌دهد. مقادیر ویژگی‌های شی و زمینه اطراف آن برای محاسبه نسبت لگاریتم احتمال هیستوگرام‌های شی و زمینه استفاده شد که پیکسل‌های شی را به مقادیر مثبت و پیکسل‌های زمینه را به مقادیر منفی نگاشت می‌دهد. به این ترتیب شی را از زمینه جدا می‌کند [۹]



شکل (۷-۲) - نتایج ردیابی در [۱۰]، تعدادی فریم از یک فیلم به همراه نقشه اطمینان قسمت مربوط به قسمت انتخاب شده توسط مستطیل

در [۱۱] نویسندگان از رویکرد دیگری مبنی بر خوشه‌بندی k میانگین ویژگی‌های رنگ و مکان استفاده کرده‌اند. ایده‌ی اصلی رویکرد به این ترتیب است که ابتدا ویژگی‌های شیء و زمینه‌ی نزدیک به شیء (که اطراف شیء قرار دارد و نه کل زمینه) را خوشه‌بندی کرده و سپس با حذف خوشه‌های مشترک، شیء را از زمینه‌ی اطرافش متمایز کرده و آن را ردیابی کرده است. در این روش نیز حضور شیء دیگری با ویژگی‌های مشابه با هدف در اطراف آن مدل هدف را ضعیف می‌کند و حتی ممکن است شیء مورد نظر را تشخیص ندهد. در [۱۲] نویسندگان ابتدا از رویکرد [۹] و ویژگی‌های رنگ برای جداسازی شیء از زمینه اطرافش استفاده کردند. در ادامه آنها از دو شبکه عصبی مجزا استفاده کردند. به یک شبکه عصبی ویژگی‌های شیء را آموزش دادند و به شبکه عصبی دیگر ویژگی‌های زمینه را آموزش دادند و در طول ردیابی آموزش شبکه‌های عصبی را گرفتن نمونه‌های جدیدی از شیء و زمینه به روز رسانی می‌کردند. از آنجا که این نوع بروز رسانی مدل شیء بصورت آهسته و با یک نرخ ثابت صورت می‌گیرد اگر تغییرات شیء بیشتر از میزان قابل تحمل باشد ردیابی دچار خطا می‌شود. همچنین جداسازی شیء از زمینه اطرافش با استفاده از نسبت هیستوگرام‌ها همانطور که گفته شد دارای حجم محاسباتی بالایی بوده و با افزایش ویژگی‌ها دچار زیادی ابعاد شده و عملاً بسیار زمانبر است. ضعف دیگر آن استفاده از تنها ویژگی رنگ است. استفاده‌ی تنها از ویژگی رنگ باعث می‌شود که ردیاب به

آسانی به اشیاء مجاور شیء اصلی که هم‌رنگ با آن هستند منحرف شود. نمونه از این وضعیت در شکل (۸-۲) نمایش داده شده است.



شکل (۸-۲) - هنگامی که زمینه رنگی نزدیک به هدف داشته باشد الگوریتم ردگیری دچار اشتباه می‌شود. در شکل بالا هدف بعد از رسیدن به نقاط سفید در زمینه اشتباه این نقاط در نظر گرفته می‌شود. [۱۲]

یک نکته‌ی قابل ذکر در رویکردهای اخیر که شیء را از زمینه جدا می‌کنند این است که سیاه‌نمای شیء می‌تواند مستقیماً یا با آستانه‌گیری از نقشه‌ی احتمال بدست آید بنابراین این روش‌ها را می‌توان زیر مجموعه‌ای از گروه ردیابی‌های سیاه‌نمای شیء نیز در نظر گرفت ولی از آنجا که این روش‌ها مستقیماً از ویژگی‌های تصویر استفاده می‌کنند و زیر مجموعه‌ای از ردیابی مبنی بر هسته است در این قسمت جای داده شده‌اند. یکی از مزیت‌های این دسته از رویکردهای ردیابی این است که با انتخاب درست ویژگی‌های مستقل از اسکلت و تغییر شکل‌های ظاهری، می‌توان اشیاء غیر صلب را نیز مدل کرد. مزیت روش‌های اخیر که جدا کردن شیء از زمینه را نیز انجام می‌دهند این است که از شیء استخراج شده می‌توان در پردازش‌های سطح بالاتر و برای تفسیر و تعبیر حرکت نیز استفاده نمود.

۲-۱-۳-روش‌های مدل‌های چندتایی

در روش‌هایی مبنی بر هسته که مدل شیء توسط قالب، هیستوگرام و غیره بیان می‌شود معمولاً مدل شیء بصورت بهنگام ساخته می‌شود بنابراین این مدل‌ها اطلاعات بدست آمده از آخرین مشاهدات شیء را بازنمایی می‌کنند. اما اشیاء ممکن است از زوایای متفاوت ظاهر متفاوتی داشته باشند و اگر زاویه‌ی مشاهده‌ی شیء در طول ردیابی تغییر شدیدی کند ممکن است ظاهر شیء متفاوت با مشاهدات آن در فریم‌های قبلی گردد و ردیابی دچار خطا گردد. برای غلبه بر این مشکل نماهای مختلف شیء از قبل می‌تواند بصورت آفلاین به یک شبکه آموزش پذیر یاد داده شود و برای ردیابی مورد استفاده قرار گیرد [۱].

به این منظور در [۱۳] نویسنده از طبقه‌بندی کننده‌های ماشین بردار پشتیبان برای ردیابی عقب وسایل نقلیه استفاده کرده است. ماشین بردار پشتیبان یک روش طبقه‌بندی کلی است که با داشتن یک مجموعه مثال‌های آموزشی مثبت و منفی بهترین ابر صفحه جداکننده بین دو کلاس را می‌یابد. ماشین‌های بردار پشتیبان را بر روی هزاران تصویر از مثال‌های مثبت (تصاویری از شیء که باید ردیابی شوند شامل تصاویر و وسایل نقلیه) و مثال‌های منفی (یعنی همه‌ی چیزهایی که نباید ردیابی شوند شامل تصاویر وسایل غیر نقلیه) آموزش داده است. این رویکرد از ردیاب و طبقه‌بندی کننده بطور متوالی استفاده می‌کند. ردیاب می‌فهمد که شیء به کجا رفته است و کلاسیفایر به آن امتیاز می‌دهد. این روش بجای کمینه کردن تفاوت شدت‌های قالب‌ها با مناطق تصویر امتیاز کلاس‌بندی ماشین بردار پشتیبان را به روی تصاویر به منظور تخمین مکان شیء پیشنهاد می‌کند. مثالی از این رویکرد به همراه امتیاز دهی ماشین بردار پشتیبان به شیء مورد ردیابی در طول دنباله در شکل (۲-۹) نمایش داده شده است. یک مزیت این رویکرد این است که اطلاعات راجع به اشیاء زمینه صریحاً در ردیاب شرکت داده می‌شوند. [۴۱] و [۴۲] نیز با رویکردی مشابه یک شبکه عصبی را بر روی مجموعه‌ای از نمونه‌های ورودی و خروجی برای ردیابی یک شیء خاص به صورت آفلاین آموزش داده‌اند. بزرگترین ضعف این

دسته از رویکردهای ردیابی، تک متغیره بودن آن است که کاربرد آن را محدود به ردیابی یک هدف از قبل تعیین شده می‌کند. ضعف دیگر این روش‌ها احتیاج آن‌ها به یک پروسه‌ی یادگیری آفلاین است.



شکل (۲-۹) - a, b, c سه فریم از دنباله ویدیوی هستند و d امتیاز ماشین بردار پشتیبان به شیء مورد ردیابی در طول ردیابی [۱۴]

بطور کلی یکی از محدودیت‌های اغلب رویکردهای مبنی بر اساس هسته و استفاده از شکل‌های هندسی ابتدایی برای بازنمای اشیاء این است که ممکن است قسمتی از شیء خارج از شکل اولیه‌ی تعریف شده قرار گیرد در حالی که قسمتی از زمینه در درون آن قرار بگیرد. این پدیده می‌تواند برای هر دو نوع جسم صلب (وقتی که وضعیت شیء عوض می‌شود) و اشیاء غیر صلب (وقتی که حرکات محلی شیء باعث تغییر در ظاهر شیء می‌گردد) مشاهده گردد. در این وضعیت مکان و حرکات شیء با بیشینه کردن شباهت مدل تخمین زده می‌شود که می‌تواند درست نباشد. برای غلبه بر این محدودیت یک

رویگرد این است که به جای اینکه شکل اولیه مستطیل یا بیضی را طوری انتخاب کنیم که کل شیء درون آن قرار گیرد هسته مستطیل یا بیضی را درون شیء قرار دهیم. رویگرد دیگر این است که شیء را توسط تابع چگالی ویژگی‌های آن مانند رنگ و بافت مدل کرده و آن را از زمینه اطرافش جدا کنیم. به عنوان رویکردی دیگر می‌توان از سیاه‌نما برای مدل کردن دقیق شیء و ردیابی آن استفاده کرد. که در ادامه به آن پرداخته می‌شود.

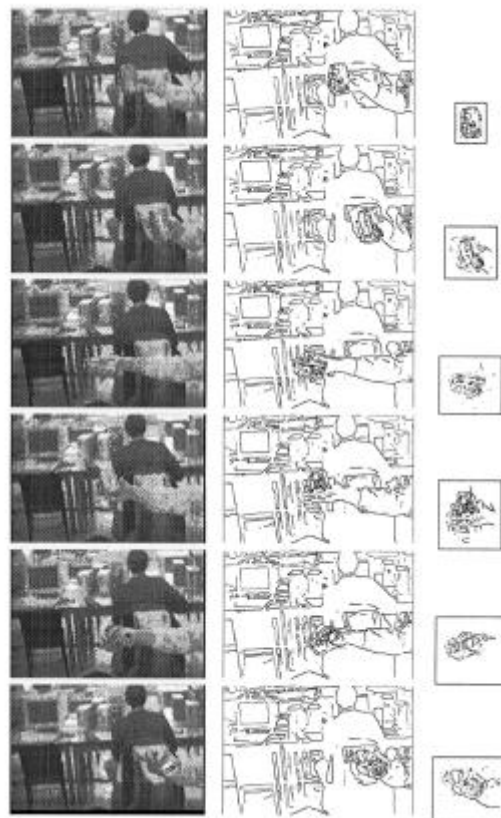
۲-۱-۳- ردیابی سیاه‌نما

اشیاء ممکن است شکل‌های پیچیده داشته باشند برای مثال دست‌ها و سر شانه‌ها و بدن انسان نمی‌تواند با شکل‌های ساده هندسی توصیف گردد در این موارد شیء را می‌توان توسط محدوده دقیق آن مدل کرد. با مشخص بودن مرزهای شیء ناحیه درون آن سیاه‌نمای شیء نامیده می‌شود. در این رویکرد اسکلت شیء می‌تواند از سیاه‌نما یا کانتور (کناره نما) شیء استخراج شود. در این روش‌ها هدف، ردیابی شیء با استفاده از لبه‌ها و مرزهای شیء است. ردیاب‌های سیاه‌نما را می‌توان به دو دسته با نام‌های روش‌های تطابق شکل و ردیابی کانتور تقسیم کرد.

۲-۱-۳-۱- روش‌های تطابق شکل

تطابق شکل‌ها می‌تواند مشابه با ردیابی بر اساس تطابق قالب انجام گیرد که سیاه‌نمای شیء و مدل وابسته‌اش در فریم فعلی مورد جستجو قرار می‌گیرد. جستجو با محاسبه شباهت شیء با مدل ساخته شده با سیاه‌نمای شیء فرض شده بر اساس فریم قبل انجام می‌گیرد. در این رویکرد فرض می‌شود که سیاه‌نما تنها از فریمی به فریم بعد جابجا می‌شود بنابراین حالت اشیاء غیر صلب در نظر گرفته نمی‌شوند. مدل شیء که معمولاً به شکل لبه‌های آن است مجدداً باز تعریف می‌شود تا تغییرات ظاهری در هر فریم پس از اینکه شیء مکان یابی شد کنترل کند. الزام است این بروز رسانی بر مشکلات ردیابی که مرتبط با تغییر زاویه‌ی دید، حرکات اجسام غیر صلب و تغییرات شدت روشنایی است غلبه کند.

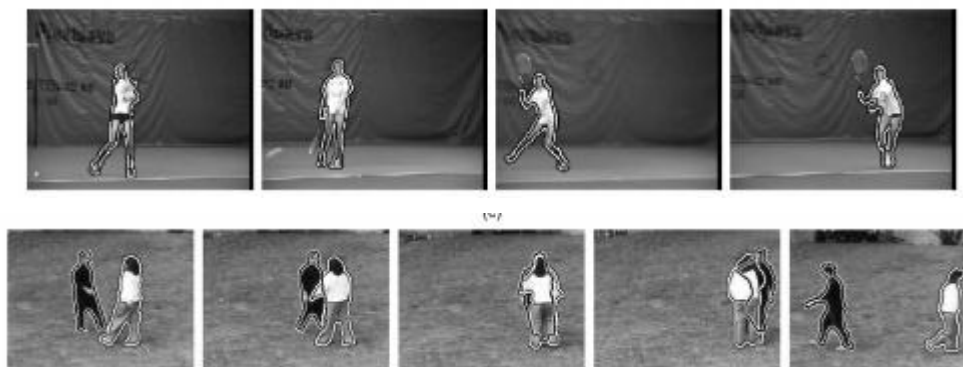
در [۱۵] تطابق شکل‌ها با استفاده از بازنمایی براساس لبه انجام گرفته است. نویسندگان از فاصله Hausdorff [۱۶] استفاده کردند تا یک سطح همبستگی بسازند بطوری که کمینه آن مکان جدید شی خواهد بود. در حوزه‌ی تطابق با استفاده از مدل لبه، معیار اندازه گیری Hausdorff بیشترین عدم تطابق را اندازه گیری می‌کند به این علت این روش بر قسمت‌هایی از طرح لبه‌ها تاکید می‌کند که توسط حرکت شی تغییر جدی نکرده باشند. برای مثال برای شخص در حال حرکت، سر و نیم تنه تغییر زیادی ندارد در حالی که حرکات دست و پا منجر به تغییرات شکل شدیدی می‌شود بطوری که حذف لبه‌های متناظر با دست و پا عملکرد را بهبود می‌دهد. مثالی از رویکرد [۱۵] در شکل (۲-۱۰) نمایش داده شده است.



شکل (۲-۱۰)-تعدادی فریم از دنباله ویدئویی، به همراه مدل هدف مورد ردیابی [۱۵]

۲-۱-۳-۲- روش‌های ردیابی کانتور

مدل‌های کانتور فعال اولین بار در سال ۱۹۸۷ ارائه شد. این روش در ابتدا از نظر دقت و سرعت با مشکلات زیادی روبرو بود ولی به تدریج توسعه یافت. کانتور فعال یک منحنی فرم‌پذیر می‌باشد که با کمینه سازی انرژی تعریف شده برای آن یک مشخصه را در تصویر که معمولاً مرز هموار یک ناحیه می‌باشد، آشکارسازی می‌نماید. در این روش کانتور اولیه در فریم قبلی بتدریج و بطور تکرار شونده در طول زمان تکامل می‌یابد. سیر کامل شدن کانتور توسط تابع انرژی هدایت می‌شود. این تکامل تدریجی نیاز دارد که بعضی قسمت‌های شیء در فریم فعلی با نواحی شیء در فریم قبلی همپوشانی داشته باشد. این عامل به همراه عواملی مانند حرکت‌های سریع هدف، پس‌زمینه شلوغ، پوشیدگی توسط زمینه، وجود لبه‌های قوی در پس‌زمینه و لرزش دوربین در حین تصویر برداری، به شدت بر روی عملکرد کانتور فعال در ردیابی تاثیر گذاشته و موجب ناموفق بودن آن شود. اغلب از اطلاعات حرکتی برای کمک به ردیابی کانتور فعال استفاده می‌شود.



شکل (۲-۱۱)- نتایج الگوریتم ردیابی کانتور [۱۷]

بطور کلی ردیابی سیاه‌نما وقتی بکار گرفته می‌شود که ردیابی کل شیء مورد نظر می‌باشد. بعضی از این الگوریتم‌ها تنها به اطلاعات مرز سیاه‌نما برای ردیابی نیاز دارد در حالیکه بقیه از کل ناحیه‌ی درون سیاه‌نما برای ردیابی استفاده می‌کنند. بطور کلی رویکردهای بر اساس ناحیه در برابر نویز مقاوم‌تر هستند و انعطاف‌پذیری بیشتری دارند. مهمترین مزیت ردیابی سیاه‌نما انعطاف آن برای

ردیابی انواع وسیعی از شکل‌های شیء است. معمولاً این دسته از روش‌ها به مساله پوشیدگی صریحاً اشاره نمی‌کنند. یک رویکرد معمول این است که فرض می‌شود حرکت ثابت یا سرعت ثابت است و در طول پوشیدگی، سیاه‌نما شیء از فریم قبل به فریم جدید تنها دارای حرکت انتقالی است. موضوع دیگر در این رویکردها قابلیت کنترل چند تیکه شدن و ادغام دوباره شیء است. برای مثال با فرض اینکه هدف مورد ردیابی فردی باشد که شیء را حمل می‌کند، وقتی که فرد شیء را رها می‌کند قسمتی از کانتور شخص روی شیء رها شده باقی می‌ماند. [۱]

۲-۱-۴- روش‌های تفریق فریم و تفریق زمینه

۲-۱-۴-۱- روش تفاضل فریم‌های متوالی

این روش نسبت به تغییرات نوری و تغییرات زائد پس زمینه مقاوم بوده و تغییرات شیء متحرک را آشکار می‌کند. ولی در استخراج همه نقاط مربوط به شیء ضعیف عمل می‌کند و معمولاً فقط نقاط مربوط به لبه‌های شیء را آشکار می‌سازد.

فرض کنید که $I_n(\mathbf{x})$ نشان دهنده شدت روشنایی فریم در نقطه $\mathbf{x}$ در فریم n ام باشد. آنگاه فریم تفاضل بصورت زیر تعریف می‌شود:

$$FD(\mathbf{x}) = I_n(\mathbf{x}) - I_{n-1}(\mathbf{x}) \quad (3-2)$$

فریم تفاضل در حقیقت تفاضل نقطه به نقطه بین دو فریم است. با فرض اینکه شرایط نوری محیط تقریباً بین دو فریم متوالی ثابت می‌باشد، مقدار فریم تفاضل در یک نقطه بیانگر تغییر در یک نقطه می‌باشد. مقدار غیرصفر تفاضل به دو دلیل ایجاد می‌شود. نخست به دلیل حرکت اشیاء موجود در تصویر، و دوم به دلیل نویز مشاهده^۱ بین مقادیر یک نقطه در دو فریم متوالی.

^۱- Observation noise

برای جداسازی حرکت هدف متحرک از نویز در فریم تفاضل از یک مقدار آستانه استفاده می‌کنند. به این صورت که اگر مقدار فریم تفاضل از یک مقدار بیشتر شد آنرا به عنوان شیء متحرک و مقدار یک را به آن پیکسل تخصیص می‌دهند در غیر این صورت آنرا به عنوان پس زمینه می‌شناسند.

$$Binary_Mask_n(\mathbf{x}) = \begin{cases} 1 & \text{if } |FD_n(\mathbf{x})| > T_n(\mathbf{x}) \\ 0 & \text{other wise} \end{cases} \quad (4-2)$$

در (4-2)، $T_n(\mathbf{x})$ ، آستانه تصمیم‌گیری برای فریم n ام در نقطه $\mathbf{x}$ می باشد. با این وجود ممکن به اشتباه نقاطی را به عنوان شیء متحرک در نظر بگیرد. این نقاط را می‌توان توسط پس پردازش حذف کرد. یک روش برای حل مشکل این است که فریم تفاضل را به بلوک‌های 4 یا 8 پیکسلی تقسیم کرد. اگر تعداد یک‌های بلوک کمتر از یک مقدار معلوم بود این گروه را بعنوان نویز در نظر گرفت. در روش دیگر، به منظور حذف این نقاط منفرد از روش تفاضل تجمعی¹ که از حافظه بیشتری در تشخیص حرکت استفاده می‌کند، استفاده می‌شود. برای این منظور فرض کنید، که $I_n(\mathbf{x})$ فریم جاری و $I_{n-1}(\mathbf{x}), I_{n-2}(\mathbf{x}), \dots, I_{n-N}(\mathbf{x})$ فریم ماقبل آن باشد. آنگاه با اعمال تفاضل بین فریم جاری و هر کدام از فریم‌های قبلی، N فریم تفاضل بدست می‌آوریم. پس از اعمال سطح آستانه به این N فریم تفاضل و بدست آوردن N فریم باینری، اگر تعداد نقاط روشن برای مختصات یک نقطه در N فریم باینری متوالی از حدی بیشتر شود، آنگاه آن نقطه بعنوان نقطه شیء در فریم باینری حاصل لحاظ خواهد شد. در حقیقت هر چند بار که یک نقطه برای تعداد بار بیشتری در N فریم باینری به عنوان شیء تلقی شود، احتمال تعلق آن نقطه به شیء بیشتر خواهد بود و احتمال وجود آن بر اثر نویز کمتر خواهد شد. ولی در عوض، تاخیر و حافظه بیشتر از مشکلات آن بشمار می‌آید. برای مثال برای دو فریم متوالی، حاصل بدین گونه محاسبه خواهد شد.

¹- Accumulative Difference

$$Binary - Mask_n(x) = \begin{cases} 1 & \text{if } (|I_n(x) - I_{n-1}(x)| > T_n(x)) \text{ and } (|I_n(x) - I_{n-2}(x)| > T_n(x)) \\ 0 & \text{otherwise} \end{cases} \quad (5-2)$$

در شکل (۱۲-۲) خروجی این روش برای آستانه ثابت نشان داده شده است. همانطور که دیده می‌شود، همه نقاط مربوط به شی آشکار نشده‌اند و فقط نقاط مربوط به لبه‌های شی متحرک بعنوان پیش‌زمینه آشکار شده‌اند. در صورتی که سرعت شی کمتر باشد، این لبه‌ها باریک‌تر می‌شوند. و هنگامی که حرکت شی متوقف شود، این روش قادر به تشخیص شی نخواهد بود.



$I(x,y,t-1)$



$I(x,y,t)$



شکل (۱۲-۲) خروجی الگوریتم تفاضل فریم های متوالی [۱۸]

۲-۴-۱-۲- روش حذف پس زمینه

همانطور که در قسمت قبل گفته شد روش تفاضل فریم‌های متوالی اطلاعات کمی در مورد شی متحرک به ما می‌دهد. در نتیجه نمی‌تواند نتایج مطلوب فرآیند ردیابی را فراهم کند.

در روش حذف پس‌زمینه، فرض می‌شود که دوربین ثابت است و ما می‌توانیم یک مدل آماری یا توزیع احتمالی برای هر نقطه از پس‌زمینه بدست‌آوریم. اغلب چنین مدلی از پس‌زمینه، نقطه‌ای می‌باشد. بدین معنی که هر نقطه به‌گونه مستقل و مجزا مدل می‌شود. یا به عبارت دیگر هر نقطه دارای تابع توزیع احتمال مخصوص به خود خواهد بود. برای هر نقطه از تصویر، مقدار روشنایی یا رنگ آن نقطه از تصویر با مدل پس‌زمینه مقایسه می‌شود و بر این اساس این مقایسه نوع نقطه اعم از پیش‌زمینه یا پس‌زمینه بودن آن، مشخص می‌گردد. بعنوان مثال، در یک روش ساده می‌توان یک تصویر میانگین را به عنوان مدل پس‌زمینه گزینش نمود و نقاطی که فاصله مقادیر روشنایی آنها از مقادیر متناظر در مدل بیشتر از آستانه بخصوصی باشد، بعنوان شی گزینش خواهند شد. مدل پس‌زمینه را می‌توان از مدل‌های آماری مختلفی برای پس‌زمینه استفاده می‌شود که در ذیل شرح داده می‌شوند:

الف- مدل میانگین^۱: برای هر نقطه (x, y) ، یک مقدار میانگین $B(x, y)$ به عنوان مقدار پس‌زمینه استفاده می‌شود. برای تعیین نوع نقطه، از لحاظ تعلق به پس‌زمینه یا شی از شرط زیر استفاده می‌کنیم. اگر رابطه ذیل (۲-۴) برقرار باشد، نقطه جزء شی است وگرنه متعلق به پس‌زمینه خواهد بود.

$$|I(x, y) - B(x, y)| > T \quad (۲-۶)$$

در رابطه فوق T مقدار آستانه تصمیم‌گیری و $I(x, y)$ مقدار روشنایی تصویر ورودی در نقطه مزبور می‌باشد. مدل پس‌زمینه باید در طول زمان برای تطبیق با تغییرات تدریجی روشنایی محیط و سایر

¹- Mean

عوامل، خود را وفق دهد. یکی از مدل‌هایی که برای تطبیق مدل پس‌زمینه در این حالت ارائه شده بصورت زیر است.

$$B_{n+1}(\mathbf{x}) = \begin{cases} \alpha B_n(\mathbf{x}) + (1-\alpha)I_n(\mathbf{x}) & \mathbf{x} \text{ is non-moving} \\ B_n(\mathbf{x}) & \mathbf{x} \text{ is moving} \end{cases} \quad (7-2)$$

در این مدل سطح آستانه نیز مورد تطبیق قرار می‌گیرد:

$$T_{n+1}(\mathbf{x}) = \begin{cases} \alpha T_n(\mathbf{x}) + (1-\alpha)(5 \times |I_n(\mathbf{x})| - B_n(\mathbf{x})) & \mathbf{x} \text{ is non-moving} \\ T_n(\mathbf{x}) & \mathbf{x} \text{ is moving} \end{cases} \quad (8-2)$$

در روابط (۵-۲) و (۶-۲)، α نرخ یا سرعت تطبیق مدل با تغییرات محیط می‌باشد. اگر $\alpha = 0$ ، اختیارشود، بیشترین تطبیق با تغییرات اعمال می‌شود و هر قدر که α به یک نزدیکتر باشد، سرعت تطبیق‌پذیری کمتر خواهد بود. از α بعنوان حافظه سیستم نیز یاد می‌شود. روابط فوق بصورت فیلتر خطی با پاسخ ضربه نامحدود^۱ می‌باشد. $B_n(\mathbf{x})$ در رابطه (۵-۲) بصورت متوسط خطی مقادیر روشنایی نقاط تصویر $I_n(\mathbf{x})$ است. در حالیکه، $T_n(\mathbf{x})$ در (۶-۲) به سمت پنج برابر انحراف معیار مقادیر $I_n(\mathbf{x})$ میل می‌کند. از این روش بعنوان الگوریتم حذف پس‌زمینه واقعی^۲ نیز یاد می‌شود. در شکل (۲-۱۳) یک نمونه از این پردازش برای مدل الف نشان داده شده است.

¹- Infinite Impulse Response (IIR)

²- Adaptive Background Subtraction



$I(x,y,t)$

$B(x,y,t)$

(الف)



(ب)

شکل (۲-۱۳) پردازش حذف پس‌زمینه (الف) تصویر رنگی جاری و پس‌زمینه. (ب) خروجی پردازش حذف پس‌زمینه [۱۹].

ب- مدل میانگین و کواریانس^۱: برای توصیف هر نقطه از تصویر از توزیع گوسی استفاده می‌شود یا به عبارت دیگر توزیع آماری هر نقطه بوسیله یک میانگین $B(x,y)$ و کواریانس آن، $C(x,y)$ توصیف خواهد شد. برای بررسی تعلق یک نقطه $I(x,y)$ ، به پس‌زمینه از فاصله ماحالانوبیس^۲ بصورت رابطه (۲-۹) استفاده خواهد شد. نقطه متعلق به پیش‌زمینه رابطه زیر را ارضاء خواهد کرد.

^۱ - Mean & Covariance

^۲ - Mahalanobis

$$[I(x, y) - B(x, y)]^T C^{-1}(x, y) [I(x, y) - B(x, y)] > T \quad (9-2)$$

ج- انحراف زمانی^۱: در این مدل برای هر نقطه مقدار ماکزیمم شدت روشنایی آن نقطه، $M(x, y)$ ، و مینیمم مقدار شدت روشنایی آن نقطه، $N(x, y)$ ، و بزرگترین قدرمطلق تفاضل مقادیر روشنایی در آن نقطه بین زوج فریم‌های متوالی نگهداری می‌شود. اگر مقدار روشنایی در این نقطه، $I(x, y)$ در تصویر ورودی شروط زیر را ارضا کند، بعنوان پیش‌زمینه تلقی می‌شود.

$$|I(x, y) - N(x, y)| > D(x, y) \quad (10-2)$$

$$|I(x, y) - M(x, y)| > D(x, y)$$

د- پیش بینی خطی^۲: مدل پیش‌بینی خطی از فیلتر وینر^۳ برای مدل کردن الگوی زمانی هر نقطه استفاده می‌کند. در این مدل از p فریم گذشته برای پیش‌بینی فریم جاری استفاده می‌شود. این مدل قادر است، تا خود را با تغییرات روشنایی وفق دهد. پس از تخمین پارامترها برای فیلتر وینر، انجام پردازش تفاضل پس‌زمینه راحت خواهد بود. مدل پیش‌بینی خطی بصورت زیر می‌باشد:

$$B_t(x, y) = - \sum_{k=1}^p a_k B_{t-k}(x, y) \quad (11-2)$$

۲-۲- حذف نویز:

در محیط بیرون^۴ با پس‌زمینه متغیر دو منبع تشخیص اشتباه^۵ وجود دارد. نخست تشخیص اشتباه برای نویز تصادفی، که در سراسر صفحه ممکن است و دوم تشخیص‌های اشتباه برای حرکات‌های کوچک در پس‌زمینه، که بوسیله مدل پس‌زمینه نمایش داده نمی‌شود. به‌عنوان مثال اگر شاخه درخت

^۱- Temporal Deviation

^۲- Linear Prediction

^۳- Wiener

^۴- Outdoor Environment

^۵- False Detection

بیشتر از آن چیزی که در مدت تشکیل مدل حرکت می‌کرد، تکان بخورد یا جابجایی‌های کوچک در محل دوربین به سبب باد به وجود آید، احتمال وقوع تشخیص‌های اشتباه زیاد خواهد بود.

مرحله دوم در تشخیص اشیاء متحرک حذف تشخیص‌های اشتباه برای حرکت‌های کوچکی است که در مدل در نظر گرفته نشده‌اند. اگر قسمتی از پس‌زمینه (یک شاخه درخت) حرکت کند و یک نقطه جدید را اشغال کند، که مقدار روشنایی جدید این نقطه، در مدل آماری پس‌زمینه نقطه وجود نداشته باشد، این نقطه به عنوان پیش‌زمینه تلقی خواهد شد. اما این نقطه به احتمال قوی متعلق به توزیع پس‌زمینه در مکان اصلی خود خواهد بود. با فرض اینکه جابجایی کوچکی ما بین دو فریم متوالی اتفاق افتاده، می‌توان با در نظر گرفتن توزیع پس‌زمینه در همسایگی نقطه مزبور، تعیین کرد که آیا این نقطه متعلق به پس‌زمینه همسایه بوده و در اثر حرکت پس‌زمینه بوجود آمده، یا واقعا نقطه ای از پیش‌زمینه است.

با اعمال آستانه در نقطه آشکار شده به عنوان شی، می‌توان بسیاری از تشخیص‌های اشتباه به علت حرکت‌های کوچک در پس‌زمینه را حذف کرد. متأسفانه با انجام این پردازش بعضی از نقاطی که به درستی نیز به عنوان شی گزینش شده بودند، نیز حذف می‌گردند. زیرا که ممکن است بعضی از نقاط آشکار شده واقعی به عنوان پیش‌زمینه در مرحله قبل، بر حسب تصادف مشابه پس‌زمینه نقاط همسایه باشند. که این مورد اغلب در مورد تصاویری که فقط با سطح روشنایی¹ نمایش داده می‌شوند، اتفاق می‌افتد. برای پرهیز از دست دادن چنین تشخیص‌های درستی، از شرط اینکه همه نقاط شی تشخیص داده شده، باید از یک همسایگی مجاور آمده باشند و نه فقط بعضی از نقاط محدود، استفاده می‌کنیم.

¹- Gray level

بدین منظور احتمال اینکه یک مؤلفه متصله^۱، از یک محل همسایگی حرکت کرده باشد، را در نظر می گیریم.

در شکل (۱۴-۲) پردازش حذف تشخیص اشتباه نشان داده شده است. در شکل (۱۴-۲ الف)، تصویر فریم جاری و در شکل (۱۴-۲ ب)، خروجی پردازش حذف پس زمینه بدون حذف تشخیص اشتباه نشان داده شده است. همانطور که دیده می شود، نقاط ریز زیادی برای حرکت شاخه های درختان در باد ایجاد شده است. در شکل (۱۴-۲ ج)، پردازش حذف تشخیص اشتباه با در نظر گرفتن مشابهت یک نقطه در همسایگی بدست آمده است. همانگونه که دیده می شود علاوه بر نقاط نویز، نقاطی از شی نیز از بین رفته اند. در شکل (۱۴-۲ د)، با در نظر گرفتن شرط حرکت همه نقاط مؤلفه متصله در همسایگی نقطه مزبور، دیگر نقاطی از شی اصلی حذف نمی شوند ولی نقاط نویز مانند حذف می شوند.



(الف)

¹ - Connected Component



(ب)



(ج)



(د)

شکل (۲-۱۴) پردازش حذف تشخیص اشتباه. الف) تصویر جاری ب) پردازش پس زمینه بدون حذف تشخیص اشتباه. ج) پردازش حذف اشتباه با در نظر گرفتن یک نقطه در همسایگی د) پردازش حذف اشتباه با در گرفتن همه مولفه متصله همسایگی [۲۰].

۲-۳-ویژگی‌های بکار رفته در ادبیات تحقیق

انتخاب ویژگی مناسب نقش اساسی در ردیابی شیء در تصویر بازی می‌کند. ویژگی‌ای مطلوب است که شیء را در فضای ویژگی از زمینه متمایز باشد. انتخاب ویژگی با مدل کردن شیء رابطه نزدیکی دارد. مثلاً برای بازنمایی‌های بر اساس هیستوگرام اغلب ویژگی رنگ استفاده می‌شود در حالی که برای بازنمایی‌های بر اساس کناره‌نما معمولاً از لبه‌های شیء به عنوان ویژگی استفاده می‌شود. بسیاری از الگوریتم‌های ردیابی نیز از ترکیبی از این ویژگی‌ها برای بهبود عملکرد ردیاب استفاده می‌کنند [۱]. در ادامه ویژگی‌های معمول در ادبیات ردیابی را بیان می‌کنیم.

۲-۳-۱-رنگ

از میان تمام ویژگی‌ها، ویژگی رنگ پر استفاده‌ترین ویژگی برای ردیابی است. اما این ویژگی جدا از محبوبیتی که دارد به تغییرات روشنایی حساس است. بنابراین در دنباله‌های که تغییرات روشنایی به ناچار در آن تاثیر گذار است از ویژگی‌های دیگر نیز به همراه رنگ برای مدل کردن شیء استفاده می‌گردد. در پردازش تصویر معمولاً از فضای رنگی R-G-B برای بازنمایی رنگ استفاده می‌شود.

۲-۳-۲-بافت

بافت، اندازه‌گیری تغییرات شدت سطح است که برای خواصی مانند همواری و نظم و ترتیب سطح، کمیت تعیین می‌کند. یک خاصیت مهم ویژگی بافت این است که نسبت به ویژگی رنگ حساسیت کمتری به تغییرات روشنایی دارد. در مقایسه با رنگ، بافت نیاز به یک مرحله پیش‌پردازشی دارد تا توصیف‌گرهای بافت را تولید کند. توصیف‌گرهای بافت متفاوتی وجود دارد. ماتریس‌های هم‌اتفاقی سطح خاکستری^۱ (GLCM) یک هیستوگرام دو بعدی است که رخدادهای همزمان مقادیر شدت

¹ Gray Level Cooccurrence Matrix

تصویر را در یک جهت و فاصله‌ی معین بیان می‌کند [۲۱]. توصیف‌گر بافت الگوی باینری محلی^۱ (LBP) با در نظر گرفتن همسایگی 3×3 در اطراف هر پیکسل و تخصیص یک عدد باینری متناسب با همسایگی‌ها به آن، بافت مورد نظر را تعیین می‌کند [۲۲]. انواع مختلف توصیف‌گرهای بافت LBP اخیراً در مرجع [۲۳] مقایسه شده است. موجک‌ها^۲ نیز از دیگر توصیف‌گرهای بافت بکار رفته در ادبیات ردیابی شی است.

۲-۳-۳-لبه

مرزهای شی معمولاً تغییر شدیدی در شدت تصویر ایجاد می‌کنند. آشکارساز لبه برای تشخیص این تغییرات استفاده می‌شود. لبه‌مهند و ویژگی بافت حساسیت کمتری نسبت به تغییرات روشنایی دارد. الگوریتم‌هایی که مرزهای اشیاء را دنبال می‌کنند معمولاً از لبه به عنوان ویژگی برای بازنمایی شی استفاده می‌کنند. آشکار ساز لبه Canny به علت سادگی و دقت محبوب‌ترین آشکارساز بکار رفته در ادبیات ردیابی شی است [۲۴].

۲-۴-پیچیدگی‌ها و موانع مساله ردیابی شی

همان گونه که بیان گردید در ردیابی شی هدف یافتن مسیر حرکت شی در طول زمان است. اگر شی بطور پیوسته قابل مشاهده باشد و شکل، اندازه و حرکات آن در طول زمان تغییر نکند در این صورت مسئله ردیابی مسئله سختی نخواهد بود. با این وجود در محیط‌های واقعی مانند یک خیابان شلوغ یا یک مرکز خرید این فرضیات ساده کننده بطور پیوسته برقرار نخواهند بود و اشیاء به خصوص افراد در حال حرکت شکل‌شان تغییر خواهد کرد بعلاوه حرکات آن‌ها ثابت نیست. تا به اینجا رویکردها و الگوریتم‌های مختلفی بیان شد و تا حدودی مزایا و مشکلات روش‌های مختلف بیان گردید مشکلات

¹ Local Binary Patterns

² wavelet

ردیابی به این عوامل محدود نمی‌شود در این قسمت یک جمع بندی از مهمترین مشکلات موجود در ردیابی دو بعدی شیء در تصاویر ارائه می‌گردد.

۲-۴-۱- کاهش ابعاد

تا کنون سیستم‌های تک دوربین و چند دوربین برای ردیابی سه بعدی اشیاء بکار گرفته شده‌اند اما این سیستم‌ها دارای مشکلات نسبتاً زیادی هستند و بنابراین همواره علاقمندی‌های زیادی برای ردیابی دو بعدی اشیاء وجود دارد. در ردیابی دو بعدی هدف ردیابی و دنبال کردن تصویر اشیائی است که جابجایی سه بعدی دارند و به صورت تبدیلات دو بعدی مدل شده‌اند. این کاهش بعد باعث از دست رفتن اطلاعات نسبتاً زیادی به علت تبدیل کردن دنیای سه بعدی به دو بعدی می‌شود.

۲-۴-۲- حرکت دوربین

متحرک بودن دوربین می‌تواند موجب تغییرات شدید و دائمی در زمینه شده و بعضی از روش‌های ردیابی شیء مانند روش تفریق پس‌زمینه را ناکارآمد کند.

۲-۴-۳- تغییرات روشنایی صحنه

در طول ردیابی روشنایی صحنه ممکن است در فریم‌های مختلف تغییر کند این تغییرات شدت روشنایی باعث تغییر رنگ شیء مورد ردیابی و زمینه شده، تشخیص و ردیابی هدف را در روش‌هایی که تنها بر اساس ویژگی رنگ عمل می‌کنند با چالش‌های جدی مواجه می‌کند.

۲-۴-۴- ساختار غیر صلب شی

شی صلب به اشیائی اطلاق می‌شود که در طول ردیابی شکل و ساختار آن‌ها تفاوتی پیدا نکند. اجسامی مانند اتومبیل‌ها از این دسته‌اند. در مقابل آن اشیاء غیر صلب اجسامی هستند که در طول ردیابی شکل آن‌ها تغییر می‌کند مانند انسان در حال حرکت. تغییر شکل ظاهری اشیاء و ساختار غیر

صلب و شکل پیچیده‌ی شیء می‌تواند ردیابی آن را با چالش مواجه کند به این علت که بعضی ویژگی‌های در نظر گرفته شده برای شیء ممکن است در طول ردیابی تغییر کند و دیگر قابل اعتماد نباشد.

۲-۴-۵- حرکات پیچیده و ناگهانی شیء

می‌توان ردیابی شیء را با در نظر گرفتن شروطی بر روی حرکت شیء ساده‌تر کرد. برای مثال اغلب الگوریتم‌های ردیابی فرض می‌کنند حرکت شیء آرام و بدون هیچ‌گونه تغییرات ناگهانی است. یک فرض دیگر برای حرکت شیء می‌تواند در نظر گرفتن سرعت ثابت یا شتاب ثابت برای شیء مورد ردیابی اساس اطلاعاتی که از قبل داریم باشد.

۲-۴-۶- تغییر شکل و اندازه شیء

از آنجایی که اشیاء می‌توانند آزادانه در محیط حرکت کنند بنابراین شکل هدف ممکن است در اثر حرکاتی مانند چرخش اندک اندک در طول ردیابی تغییر کند. هر چه تعداد فریم‌های مورد ردیابی ما افزایش یابد ظاهر شیء ممکن است دچار تغییرات بزرگتری شود. این تغییرات باید توسط الگوریتم ردیابی تشخیص داده شده و مدل هدف تنظیم گردد.

۲-۴-۷- پوشیدگی شیء

پوشیدگی یا انسداد به وضعیتی اطلاق می‌شود که قسمتی یا تمامی شیء مورد ردیابی توسط شیء دیگری که در صحنه وجود دارد پوشانده شود و شیء مورد ردیابی پشت آن قرار گیرد. اگر زمینه دارای ویژگی‌های مشابه با شیء باشد (یعنی هر چه پس زمینه به شیء شبیه‌تر باشد) جداسازی و تشخیص شیء مشکل‌تر خواهد بود و ردیابی شیء با چالش مواجه می‌شود. برای مثال در روش‌های ردیابی که بر اساس ویژگی رنگ عمل می‌کنند اگر زمینه و شیء دارای رنگ مشابه باشند ردیابی با مشکل مواجه می‌شود.

۲-۴-۸-التزامات پردازش به‌هنگام

بعضی روش‌های ردیابی شیء بر رویکردهای غیر علی^۱ تکیه دارند و برای ردیابی شیء در فریم نیاز به فریم‌های آتی دارند (مانند بعضی از انواع روش‌های ردیابی بر اساس تفریق پس‌زمینه که به تمام فریم‌ها برای ساخت پس‌زمینه نیاز دارند). این روش‌ها ذاتاً قابلیت پیاده‌سازی به منظور ردیابی به‌هنگام اشیاء را ندارند. علاوه بر این، یک سیستم ردیاب شیء برای اینکه بتواند به‌هنگام تصاویر ویدئویی آنلاین را انجام می‌دهد باید قادر باشد هر فریم را در مدت کوتاهی پردازش کند. این سرعت پردازش به عواملی مانند سرعت حرکت شیء و اندازه شیء وابسته است. برای ردیابی یک حرکت روان از دید انسان سرعت پردازش حداقل ۱۵ فریم در ثانیه مورد نیاز است. بنابراین برای این منظور به یک الگوریتم با عملکرد بالا همراه با پیاده‌سازی بهینه شده احتیاج است.

۲-۵-جمع بندی

در این فصل ابتدا به منظور فراهم آوردن یک دید کلی از مساله و راه حل‌های ارائه شده تاکنون، یک طبقه‌بندی از روش‌های ردیابی متداول ارائه شد و نقاط ضعف و قوت آن‌ها بیان گردید. روش‌های پیشنهادی این پایان‌نامه جزو روش‌های ردیابی براساس ویژگی از زیر شاخه‌ی ردیابی بر اساس هسته است که روش‌های ردیابی براساس آن با جزئیات بیشتری بررسی گردید. پس از معرفی روش‌های پیشین، متداول‌ترین ویژگی‌های بکار رفته در ادبیات ردیابی شیء معرفی شد و در پایان این فصل پیچیدگی‌ها و موانع مساله‌ی ردیابی شیء بیان گردید. در فصل بعد روش پیشنهادی که زیر مجموعه‌ای از روش‌های ردیابی بر اساس ویژگی است ارائه خواهد شد.

¹ Non-causal (Acasual)

فصل سوم

ماشین بردار پشتیبان

مقدمه

طبقه‌بندی را می‌توان به صورت ساده، قرار دادن داده‌ای مانند x در کلاس مربوط به خود تعریف نمود. چنین فرآیندی به دو صورت نظارت شده و نظارت نشده صورت می‌پذیرد که در اکثر موارد طبقه‌بندی نظارت شده بر غیر نظارت شده ارجحیت دارد [۲۵].

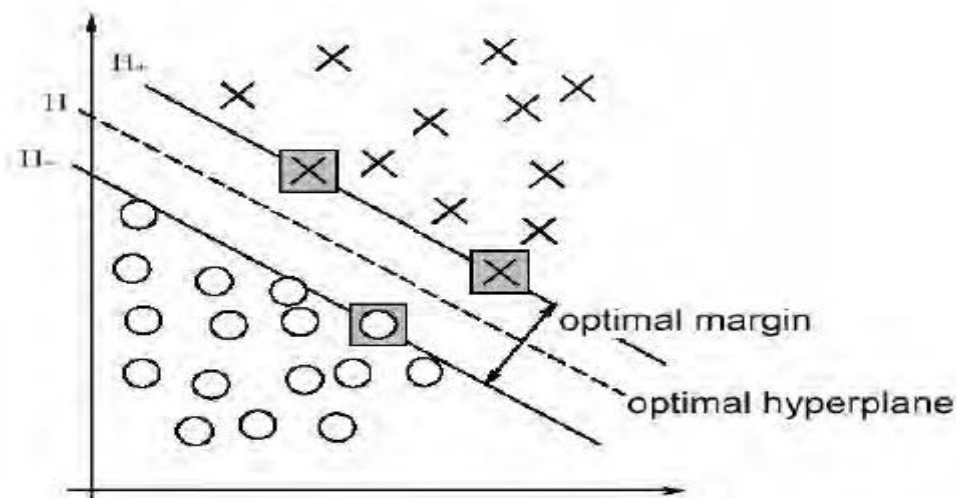
همان‌گونه که بیان گردید، الگوریتم‌های متفاوتی در فرآیند طبقه‌بندی داده‌ها مورد استفاده قرار می‌گیرند که دارای نقاط قدرت و ضعف خاصی می‌باشند. در این فصل به معرفی اصول و مبانی ماشین برداری پشتیبان که تا حدود زیادی توانسته ضعف‌های موجود در سایر الگوریتم‌ها را پوشش دهد به طور کامل مورد بحث قرار خواهد گرفت.

۳-۱- ابر صفحه جداساز

ماشین بردار پشتیبان می‌تواند به خوبی در مسائل طبقه‌بندی دو کلاسه یا چند کلاسه استفاده گردد. برای سادگی درک قضیه از مسائل طبقه‌بندی دو کلاسه استفاده می‌کنیم. هدف از این گونه مسائل ایجاد یک معیار طبقه‌بندی است که برای نمونه‌های آزمایشی به خوبی قابل استفاده باشد و در عین حال از قابلیت تعمیم خوبی برخوردار باشد. به عنوان مثال، شکل (۳-۱) را در نظر گرفته بگیرید. همان‌گونه که مشاهده می‌شود برای داده‌های شکل (۳-۱) صفحات جداساز خطی زیادی وجود دارد که قادر به جداسازی داده‌ها می‌باشند اما تنها یکی از آن‌ها دارای بیشترین حاشیه^۱ (فاصله بین صفحه جداساز و نزدیکترین نقاط) خواهد بود. این طبقه‌بندی کننده خطی، ابر صفحه جداساز^۲ بهینه نامیده می‌شود که انتظار می‌رود بتواند مرز به دست آمده را به تمام محدوده‌های ممکن تعمیم دهد. ابر صفحه جداساز بهینه در شکل (۳-۱) مشاهده می‌شود. [۲۶]

^۱ - Margin

^۲ - Hyperplane



شکل (۱-۳): صفحه جداساز بهینه با حداکثر مقدار حاشیه [۲۶]

۳-۲- صفحه بهینه برای الگوهای خطی تفکیک پذیر

در این گونه از موارد فرض می‌شود مسئله‌ای برای جداسازی مجموعه نمونه‌های آموزشی که متعلق به دو کلاس جداگانه هستند، وجود دارد به گونه‌ای که

$$D = \{x_i, y_i\} \quad \text{where} \quad i = 1 \dots N, y_i \in \{1, -1\}, x \in R^n \quad (1-3)$$

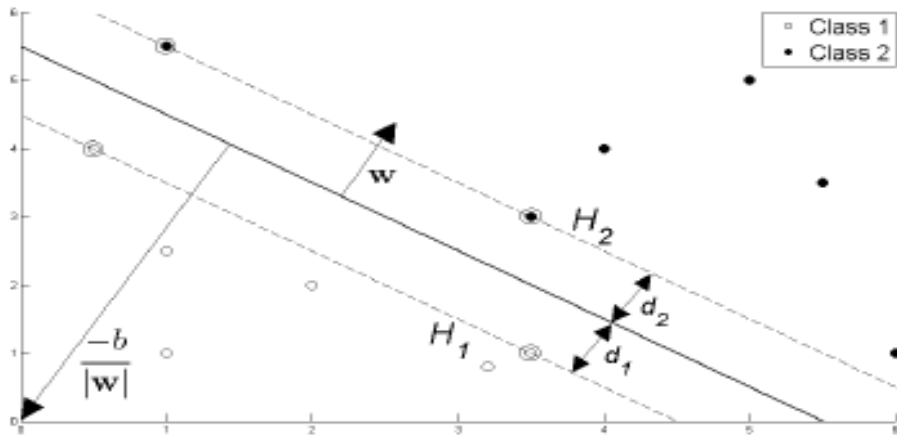
در اینجا فرض می‌کنیم داده‌ها بطور خطی جدا پذیر هستند، به این معنی که می‌توانیم آنها را با یک فوق صفحه از یکدیگر جدا کنیم. در این حالت برای صفحه جداساز بهینه معادله (۲-۳) در نظر گرفته می‌شود.

$$w^T \cdot x + b = 0 \quad (2-3)$$

W : بردار عمود بر ابرصفحه است.

$\frac{b}{\|w\|}$ فاصله عمود از ابرصفحه تا مبدا است.

بردارهای پشتیبان، نزدیکترین داده‌ها به ابرصفحه‌اند. و هدف ماشین بردار پشتیبان یافتن ابرصفحه‌ای است که بیشترین فاصله را از این داده‌ها داشته باشد.



شکل (۳-۲): صفحه جداساز و حاشیه‌ها [۲۷]

با توجه به شکل مقادیر w, b را به گونه‌ای انتخاب می‌کنیم که شرایط زیر را برقرار کند.

$$w^T \cdot x + b \begin{cases} \geq 1 & \text{for } y_i = 1 \\ \leq -1 & \text{for } y_i = -1 \end{cases} \quad (۳-۳)$$

می‌توان معادله بالا را بصورت زیر با هم ترکیب کرد.

$$y_i(x_i \cdot w + b) \geq 1 \quad (۴-۳)$$

معادله دو صفحه‌ای که بردارهای پشتیبان بر روی آنها قرار دارند بصورت زیر می‌باشد.

$$w^T \cdot x + b = 1 \quad \text{for } H_1 \quad (۵-۳)$$

$$w^T \cdot x + b = -1 \quad \text{for } H_2 \quad (۶-۳)$$

d_1, d_2 را بترتیب فاصله H_1, H_2 از ابرصفحه در نظر می‌گیریم. فاصله H_1, H_2 از ابرصفحه برابر

است. یعنی $d_1 = d_2$ هدف ماکزیمم کردن حاشیه است.

از شکل می‌توان نتیجه گرفت که فاصله میان H_1, H_2 برابر با $\frac{2}{\|w\|}$ است. ماکزیمم کردن حاشیه معادل

مینیمم کردن $\|w\|$ است. بنابراین هدف، حل معادله زیر است.

$$\min \|w\| \quad (7-3)$$

$$s. t \quad y_i(x_i \cdot w + b) \geq 1$$

مینیمم کردن $\|w\|$ معادل مینیمم کردن $\frac{1}{2}\|w\|^2$ است پس معادله (۶-۳) به صورت زیر می‌شود.

$$\min \frac{1}{2} \|w\|^2 \quad (8-3)$$

$$s. t \quad y_i(x_i \cdot w + b) \geq 1$$

که این یک مسئله بهینه سازی برنامه ریزی غیر خطی است. برای حل آن از ضرایب لاگرانژ

α که $\alpha_i \geq 0 \forall i$ استفاده می‌کنیم.

$$L_p = \frac{1}{2} w^T \cdot w - \sum_{i=1}^N \alpha_i [y_i(x_i \cdot w + b) - 1] \quad (9-3)$$

باید L_p نسبت به w, b مینیمم و نسبت به α_i ماکزیمم گردد. برای این کار نسبت به w و b مشتق جزئی گرفته شده و مساوی صفر قرار داده شود.

$$\frac{\partial L_p}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^N \alpha_i y_i x_i \quad (10-3)$$

$$\frac{\partial L_p}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (11-3)$$

و با جایگزاری L_p در به معادله زیر می‌رسیم. که باید نسبت به α ماکزیمم شود.

$$L_D = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (12-3)$$

$$s. t \quad \alpha_i \geq 0 \quad \sum_{i=1}^N \alpha_i y_i = 0$$

معادله (۱۲-۳) دوگان معادله (۹-۳) می‌باشد، با حل آن مقادیر α_i بدست می‌آید.

$$w_o = \sum_{i=1}^N \alpha_i y_i x_i \quad (13-3)$$

w_o : بردار عمود بر ابرصفحه بهینه

داده‌های متناظر با $\alpha_i > 0$ بردارهای پشتیبان (x_s) می‌باشند و از قبل می‌دانیم که پشتیبان‌ها روی ابرصفحه‌های حاشیه می‌افتند. بنابراین مقدار b از رابطه زیر بدست می‌آید.

$$b_o = \frac{1}{N_s} \sum_{s=1}^{N_s} (y_s - \sum_{m=1}^{N_s} \alpha_m y_m x_m \cdot x_s) \quad (14-3)$$

b_o : عرض از مبدا ابرصفحه بهینه

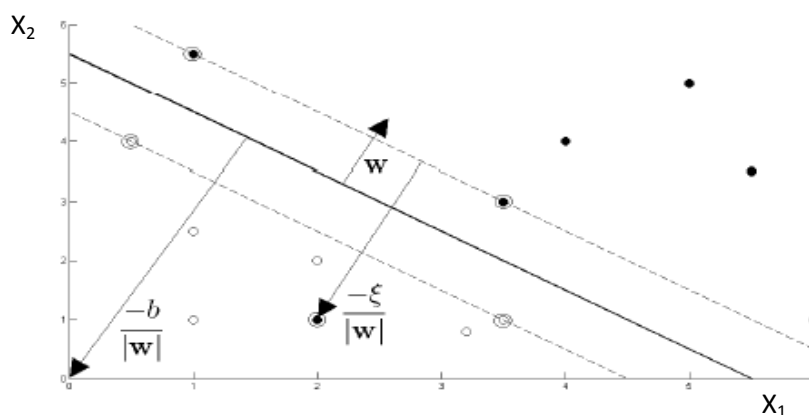
N_s : تعداد بردارهای پشتیبان

با بدست آوردن مقادیر w_o و b_o معادله صفحه جداساز بصورت زیر می‌شود. [۳۲]

$$w_o^T \cdot x + b_o = 0 \quad (15-3)$$

۳-۳- صفحه بهینه برای الگوهای خطی تفکیک‌ناپذیر

گاهی در سیستم‌های خطی شرایطی ایجاد می‌شود که تعدادی از داده‌ها در کلاس مربوط به خود قرار نمی‌گیرند (شکل (۳-۳)).



شکل (۳-۳): سیستم‌های خطی جداناپذیر با میزان خطای ξ_i [۲۷]

در این صورت ξ_i میزان خطای طبقه‌بندی را معرفی می‌کنیم و تابع هدف و محدودیت‌ها بصورت زیر تغییر می‌کنند.

$$\begin{aligned} \text{Min}_{w,b} &= \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \\ \text{S.t} \quad & y_i (w^T x + b) \geq 1 - \xi_i \end{aligned} \quad (۱۶-۳)$$

$$\xi_i \geq 0$$

در رابطه فوق، C ضریب موازنه^۱ جهت بیشینه‌نمودن حاشیه‌ها و کمینه‌سازی خطای تابع است. همان‌گونه که پیش از این نیز بیان گردید، در شرایطی مثل این از ضرایب لاگرانژ استفاده می‌شود که در رابطه اخیر با در نظر گرفتن β, α به عنوان ضرایب لاگرانژ، معادله (۱۶-۳) به صورت رابطه

(۱۷-۳) بیان می‌شود .

$$L_p(w, b, \xi, \alpha, \beta) = \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i \{ y_i [w^T x_i + b] - 1 + \xi_i \} - \sum_{i=1}^N \beta_i \xi_i \quad (۱۷-۳)$$

لاگرانژ کلاسیک دوگانه قادر است مسئله اولیه معادله (۱۷-۳) را به مسئله دوگانه آن تبدیل کند. مسئله دو گانه این رابطه، توسط معادله (۱۸-۳) تعریف می‌شود.

$$\text{Max} W(\alpha, \beta) = \text{Max}_{\alpha, \beta} (\text{Min}_{w, b, \xi} L(w, b, \alpha, \xi, \beta)) \quad (۱۸-۳)$$

اگر از رابطه (۱۸-۳) نسبت به w, b و ξ_i مشتق گرفته شده و مساوی صفر قرار داده شود، مقادیر معادله (۱۹-۳) به دست می‌آیند.

¹ -Trade-Off

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^N \alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (19-3)$$

$$\frac{\partial L}{\partial \xi} = 0 \Rightarrow \alpha_i + \beta_i = C$$

با قرار دادن این روابط در معادله (۱۸-۳)، معادله اساسی ماشین بردار پشتیبان در حالت خطی جداناپذیر به دست می‌آید که به صورت رابطه (۲۰-۳) خواهد بود.

$$\begin{aligned} \text{Max } L_d(\alpha) &= \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N y_i y_j \alpha_i \alpha_j x_i^T x_j \\ \text{S.t } &\begin{cases} 0 \leq \alpha_i \leq C \\ \sum_{i=1}^N \alpha_i y_i = 0 \end{cases} \quad (20-3) \end{aligned}$$

همان‌گونه که مشاهده می‌شود، تابع هدف سیستم‌های جداناپذیر خطی مشابه با سیستم‌های جداپذیر خطی است و تنها تفاوت این دو، اصلاح کران‌های ضرایب لاگرانژ می‌باشد. پارامتر C نیز که قابلیت کنترل ظرفیت اضافی طبقه‌بندی کننده را مشخص می‌سازد، در این سیستم‌ها باید تعیین گردد. [۲۷]

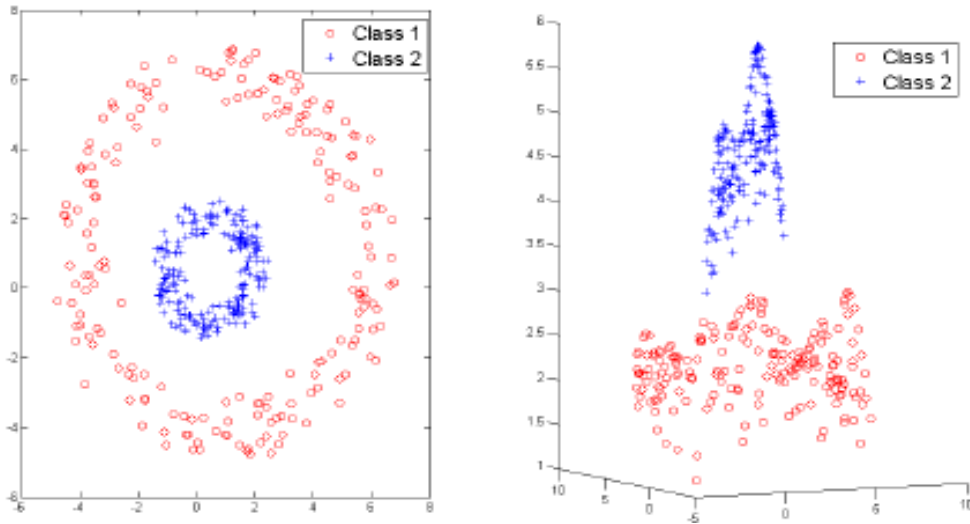
۳-۴- ماشین برداری پشتیبان برای الگوهای غیرخطی

در بخش‌های قبلی ماشین بردار پشتیبان برای حالت خطی در نظر گرفته شد. در صورتی که در بسیاری از مسائل الگوها بصورت خطی جداپذیر نیست. سوالی که مطرح می‌شود این است که آیا ماشین بردار پشتیبان برای این دسته نیز قابل تعمیم است.

مطالعات انجام گرفته در این زمینه منجر به ارائه تئوری مرسر در حل این‌گونه از مسائل شد. ایده‌ی اصلی این تئوری بصورت زیر است.

(۱) نگاشت غیرخطی از فضای بردار ورودی به فضای ویژگی با بعد بالاتر

(۲) بدست آوردن ابر صفحه بهینه برای ویژگی‌های جداسازی در مرحله ۱



شکل (۳-۴): نگاشت غیر خطی از فضای ورودی به فضای ویژگی [۲۷]

۳-۴-۱ ضرب داخلی کرنل

فرض کنید x یک بردار از فضای ورودی با بعد n باشد و $\{\varphi_j(x)\}_{j=1}^{n_1}$ یک انتقال غیرخطی از فضای

ورودی به فضای ویژگی باشد. n_1 بعد فضای ویژگی است. فرض کنید $\varphi_i(x)$ قبلاً تعریف شده است.

ابر صفحه در فضای ویژگی بصورت زیر خواهد بود.

$$\sum_{j=1}^{n_1} w_j \varphi_j(x) + b = 0 \quad (۳-۲۱)$$

یک مجموعه از وزن‌های خطی ارتباط دهنده بین فضای ویژگی و فضای خروجی $\{w_j\}_{j=1}^{n_1}$:

رابطه (۳-۲۱) را می‌توان بصورت زیر خلاصه کرد.

$$\sum_{j=0}^{n_1} w_j \varphi_j(x) = 0 \quad (22-3)$$

در این صورت $\varphi_0(x)=1$, $w_0=b$ خواهد بود.

در حالت برداری داریم

$$w^T \cdot \varphi(x) = 0 \quad (23-3)$$

$$\varphi(x) = [\varphi_0(x) \quad \cdots \quad \varphi_{n_1}(x)]^T, \quad \varphi_0(x) = 1$$

با توجه به اینکه داده‌ها در فضای ویژگی تفکیک پذیر هستند پس داریم

$$w = \sum_{i=1}^N \alpha_i y_i \varphi(x_i) \quad (24-3)$$

$\varphi_i(x)$: بردار ویژگی مربوط به نمونه‌ی x_i

با جایگزاری معادله (24-3) در (23-3) داریم

$$\sum_{i=1}^N \alpha_i y_i \varphi^T(x_i) \varphi(x) = 0 \quad (25-3)$$

ضرب داخلی کرنل را بصورت زیر تعریف می‌کنیم.

$$K(x, x_i) \triangleq \varphi^T(x_i) \varphi(x) = \sum_{j=0}^{n_1} \varphi_j(x) \varphi_j(x_i) \quad i = 1 \dots N \quad (26-3)$$

با این تعریف دیده می‌شود که ضرب داخلی کرنل یک تابع متقارن از متغیرهایش است. یعنی

$$K(x, x_i) = K(x_i, x)$$

با استفاده از ضرب داخلی کرنل می‌توانیم یک فوق صفحه در فضای ویژگی بسازیم بدون اینکه آنرا در

فضای ویژگی بطور صریح بررسی کنیم.

در نتیجه با جایگزاری (۳-۲۶) در (۳-۲۵) فوق صفحه بهینه به صورت زیر تعریف می‌شود.

$$\sum_{i=1}^N \alpha_i y_i K(x, x_i) = 0 \quad (۳-۲۷)$$

۳-۴-۲-نظریه مرسر:

فرض کنید $K(x, x')$ یک کرنل پیوسته متقارن که در بازه‌ی بسته $x, x' \in [a, b]$ تعریف شده است. کرنل $K(x, x')$ بصورت زیر بسط داده می‌شود.

$$K(x, x') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(x) \varphi_i(x') \quad , \quad \lambda_i > 0 \quad (۳-۲۸)$$

$\varphi_i(x)$: تابع ویژه

λ_i : مقدار ویژه

برای اینکه این بسط موجود و کاملاً همگرا و یکنواخت باشد. شرط لازم و کافی است که

$$\int_a^b \int_a^b K(x, x') \psi(x) \psi(x') dx dx' \geq 0 \quad (۳-۲۹)$$

برای هر $\psi(\cdot)$ که در رابطه روبرو صدق کند برقرار باشد. $\int_a^b \psi^2(x) dx < \infty$

با توجه به اینکه $\lambda_i > 0$ در نتیجه کرنل $K(x, x')$ مثبت معین است.

۳-۴-۳- طراحی بهینه ماشین بردار پشتیبان:

معادله (۳-۲۶) به ما اجازه می‌دهد یک رویه^۱ تصمیم‌گیری غیر خطی در فضای ورودی داشته باشیم (این رویه در فضای ویژگی خطی است).

با داده‌های آموزشی (x_i, y_i) where $i=1, 2, \dots, N$ ضرایب لاگرانژ را طوری پیدا می‌کنیم که تابع هدف زیر ماکزیمم شود.

$$L_D = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j K(x, x_j) \quad (3-30)$$

$$s.t \quad 0 \leq \alpha_i \leq C \quad , \quad \sum_{i=1}^N \alpha_i y_i = 0$$

بعد از بدست آوردن α_i ها داریم.

$$w_o = \sum_{i=1}^N \alpha_i y_i \varphi(x_i) \quad (3-31)$$

اولین عنصر w_o مقدار بایاس است. و یا معادله صفحه جدا کننده بصورت معادله (۳-۲۷) می‌شود.

در جدول (۳-۱) چند کرنل آورده شده است. [۲۸]

¹ surface

جدول (۳-۱) خلاصه از ضرب داخلی کرنل‌ها [۲۸]

Type of support vector machine	Inner product kernel $K(\mathbf{x}, \mathbf{x}_i), i = 1, 2, \dots, N$	Comments
Polynomial learning machine	$(\mathbf{x}^T \mathbf{x}_i + 1)^p$	Power p is specified <i>a priori</i> by the user
Radial-basis function network	$\exp\left(-\frac{1}{2\sigma^2} \ \mathbf{x} - \mathbf{x}_i\ ^2\right)$	The width σ^2 , common to all the kernels, is specified <i>a priori</i> by the user
Two-layer perceptron	$\tanh(\beta_0 \mathbf{x}^T \mathbf{x}_i + \beta_1)$	Mercer's theorem is satisfied only for some values of β_0 and β_1

۳-۵- نقاط قوت و ضعف ماشین بردار پشتیبان:

نقاط قوت ماشین بردار پشتیبان که باعث برتری این ماشین نسبت به سایر روش‌های طبقه‌بندی شده است را می‌توان به صورت زیر بیان کرد.

۱- فرآیند آموزش این ماشین بسیار آسان است (نسبت به شبکه‌های عصبی دیگر و سیستم فازی) و مانند دیگر شبکه‌های عصبی در نقطه بهینه محلی گیر نمی‌کند.

۲- در انتقال داده‌ها به فضای بالاتر به خوبی عمل نموده و عملکرد خوبی در سیستم‌های غیرخطی دارد.

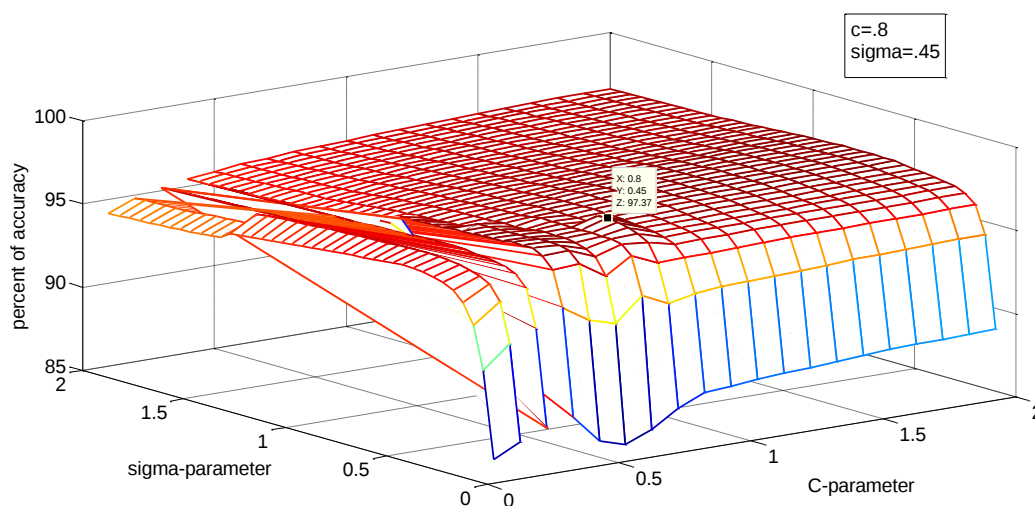
۳- موازنه بین پیچیدگی طبقه‌بندی کننده و میزان خطا قابل کنترل است.

تنها مشکل ماشین برداری که به عنوان نقطه ضعف این ماشین بیان شده، نیاز این ماشین به انتخاب کرنل مناسب است زیرا در صورت عدم مناسب بودن آن، نتایج ارائه شده رضایت بخش نخواهند بود.

۴- نیاز به تعداد نمونه‌های کم آموزشی که از جمله نقاط قوت اصلی این روش به حساب می‌آید.

۳-۶- یافتن مقادیر بهینه‌ی سیگما و c برای ماشین بردار پشتیبان:

به منظور کلاس‌بندی بهتر داده‌ها توسط ماشین بردار پشتیبان لازم است که پارامترهای قابل تنظیم آن را طوری بیابیم که کمترین خطا را برای داده‌های آزمایشی داشته باشد. برای این منظور در به ازای مقادیر مختلف سیگما و c ماشین بردار پشتیبان را آموزش می‌دهیم (C و سیگما در [۰.۱ ۲] در نظر گرفته شدند)، و سپس به ازای مقادیر مختلف درصد صحت کلاس‌بندی را توسط یک مجموعه که از مجموعه داده‌ی استخراج شده از عکس‌های منتخب، بررسی می‌کنیم [۲۹]. داده‌های آموزشی را یک مجموعه ۳۰۰ تایی از هر کلاس و مجموعه آزمایشی را یک مجموعه ۲۰۰۰ تایی از هر کلاس در نظر می‌گیریم. در شکل زیر نتایج حاصل نشان داده شده است. همان طور که مشاهده می‌شود به ازای $C=0.8$, $\sigma=0.45$ بهترین درصد دقت حاصل شده است. همچنین می‌توان برای بدست آوردن ضرایب بهینه ماشین بردار پشتیبان از الگوریتم‌های هوشمند مانند PSO، الگوریتم ژنتیک، الگوریتم کلونی مورچگان و غیره استفاده کرد.



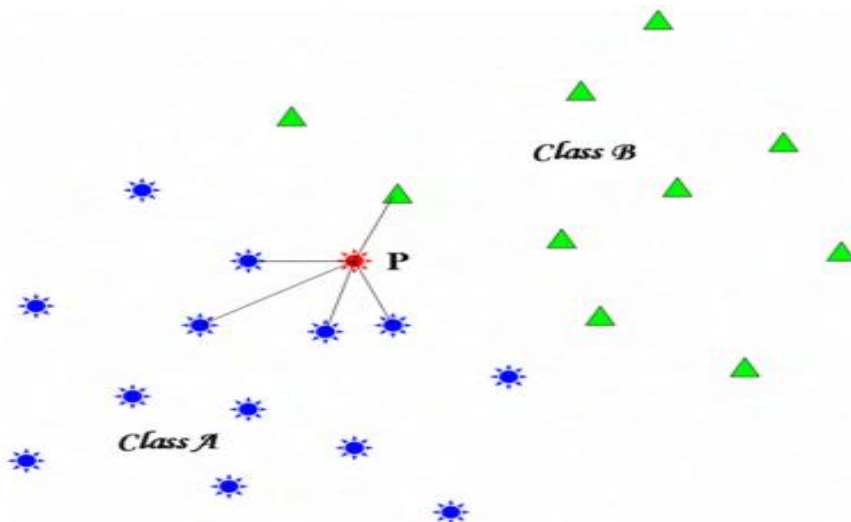
شکل (۳-۵) - نتایج حاصل برای انتخاب بهینه پارامترهای ماشین بردار پشتیبان

۳-۷- ماشین بردار پشتیبان برای مجموعه داده‌های خیلی بزرگ:

رابطه‌ی زمان آموزش در ماشین بردار پشتیبان استاندارد تابعی درجه سه از تعداد داده‌هاست. بنابراین هنگامی که داده‌های خیلی زیاد هستند نمی‌توان از ماشین بردار پشتیبان استاندارد استفاده کرد. طبقه‌بند ماشین بردار پشتیبان فقط به داده‌های نزدیک مرز جداکننده بستگی دارد. بنابراین نیاز داریم تا این داده‌ها را حفظ و بقیه داده‌ها را حذف کنیم. در زیر روش انتخاب داده‌های موثر برای ماشین بردار پشتیبان توضیح داده شده است.

فرض کنید داده‌ها همانند پیکسل‌های عکس در نظر گرفته شوند و هر کلاس رنگ مخصوص به خود را داشته باشد. مرز تصمیم‌گیری همانند لبه در تصویر می‌تواند در نظر گرفته شود. با توجه به نظریه ماشین بردار پشتیبان نمونه‌های نزدیک به مرز جداکننده بیشترین احتمال را برای اینکه بردار پشتیبان انتخاب شوند را دارند. بنابراین نمونه‌های نزدیک مرز تصمیم‌گیری می‌توانند بدین جهت همانند نقاط لبه انتخاب شوند.

آشکارسازی لبه یک اصطلاح در پردازش تصویر و بینایی ماشین است. توصیف لبه یکی از مسائل اصولی در پردازش تصویر است. به طور مشابه در مساله کلاس‌بندی نمونه‌های نزدیک به مرز تصمیم‌گیری بسیار مهم هستند. این نمونه‌ها بیشترین احتمال برای انتخاب بردار پشتیبان هستند. بنابراین نیاز است تا آن‌ها را در طی فرایند کاهش داده‌ها حفظ کنیم. در پردازش تصویر لبه‌ها بر طبق این اصل که تغییرات روشنایی در لبه‌ها شدید است، پیدا می‌شوند. آشکارسازی در مسئله کلاس‌بندی از این هم آسان‌تر است. [۳۰]



شکل (۳-۶)- آشکارسازی لبه در مسئله کلاس‌بندی [۳۰]

در پردازش تصویر لازم است تا تمام پیکسل‌های همسایگی را برای تعیین تغییرات شدت روشنایی در رنگ در نظر بگیریم. ولی در این مدل آشکارسازی لبه، این قانون به m تا از نزدیکترین نمونه‌ها تغییر می‌کند. در این پایان‌نامه فرض می‌کنیم که $m=5$ باشد. برای یک نمونه داده‌ی مشخص P اگر حداقل یکی از نمونه‌های همسایه دارای کلاس متفاوت بود آن داده را به عنوان لبه ذخیره می‌کنیم. در بعضی از موارد استفاده کردن به تنهایی از داده‌های لبه منجر به *overfitting* در ماشین بردار پشتیبان می‌شود. از این رو نیاز است تا تعدادی داده که ویژگی‌های ساختاری همه داده‌ها را بیان می‌کنند را اضافه کنیم. برای این کار از فنون خوشه‌بندی برای هر کلاس استفاده می‌کنیم. که برای این کار از روش *k_mean* استفاده می‌کنیم.

نهایتاً روش کار را می‌توان به صورت زیر خلاصه کرد.

مرحله ۱- انتخاب m

مرحله ۲- استفاده از آشکارسازی لبه تا نمونه‌های اطراف لبه را انتخاب کنیم.

- یافتن m تا نمونه‌های اطراف هر نمونه در مجموعه داده‌ها

- بررسی کردن کلاس داده‌ها: اگر حداقل یکی از همسایگی‌ها در کلاس متفاوتی از نمونه

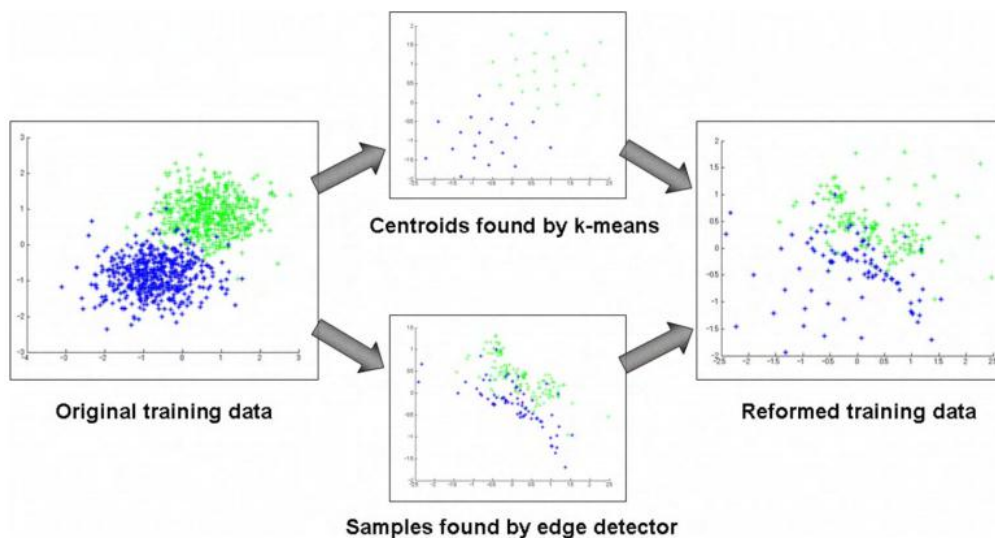
آزمایشی بود ، نمونه آزمایشی را حفظ می‌کنیم. در غیر این صورت آن را حذف می‌کنیم.

مرحله ۳- تعیین مقدار k و انجام k_mean تا مراکز هر خوشه برای تمام داده‌ها مشخص شود.

مرحله ۴- بازسازی داده‌های آموزشی با استفاده از نمونه‌های انتخاب شده از آشکارسازی لبه و مراکز

خوشه‌ها که توسط k_mean بدست آمده‌اند.

در شکل (۷-۳) این روش نشان داده شده است.

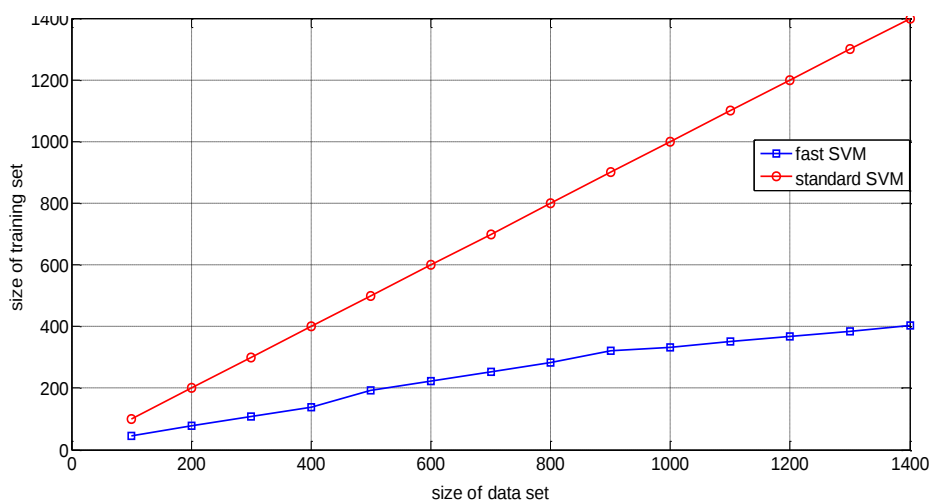


شکل (۷-۳) - کاهش داده‌های آموزشی [۳۰]

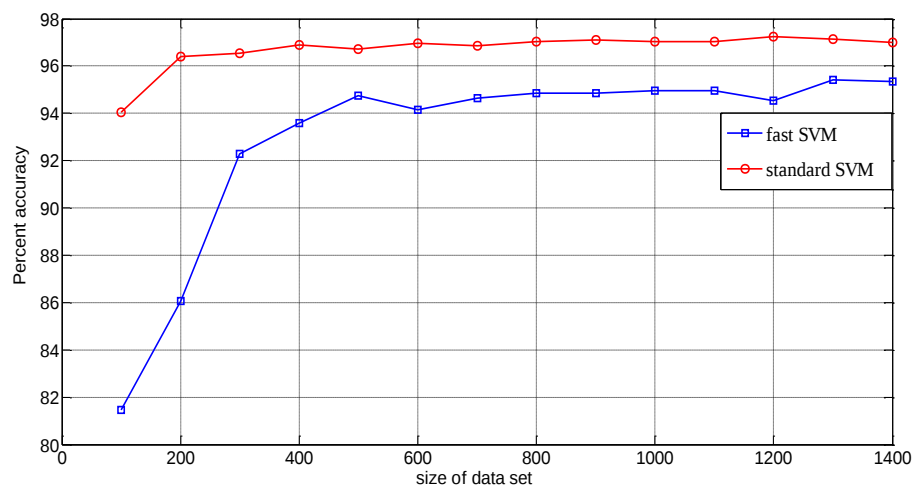
در شکل‌های زیر تفاوت روش استاندارد ماشین بردار پشتیبان و روش ارائه شده در بالا را مشاهده

می‌کنیم. مشاهده می‌شود که درصد خطا تنها مقدار اندکی افت کرده ولی زمان لازم برای آموزش به

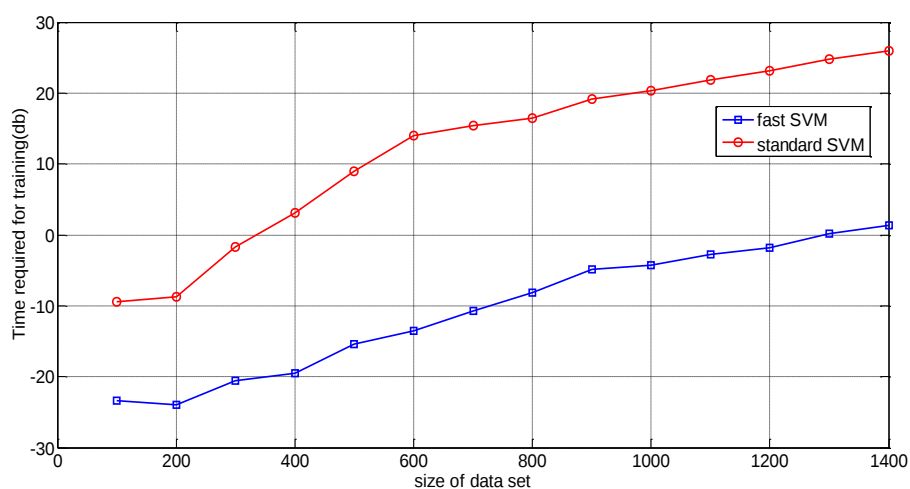
مقدار قابل ملاحظه‌ای کاهش یافته است.



شکل (۳-۸) - داده‌های آموزشی در ماشین بردار پشتیبان استاندارد و روش ارائه شده



شکل (۳-۹) - مقایسه درصد درستی ماشین بردار پشتیبان استاندارد و روش ارائه شده



شکل (۳-۱۰) - مقایسه زمان لازم جهت آموزش ماشین بردار پشتیبان استاندارد و روش ارائه شده

فصل چهارم

روش پیشنهادی و جزئیات پیاده‌سازی

مقدمه

در این فصل ابتدا الگوریتم آشکارسازی و ردگیری را بطور مختصر توضیح می‌دهیم. سپس به طور ویژه هر یک از قسمت‌های مختلف را توضیح خواهیم داد.

۴-۱- آموزش ماشین بردار پشتیبان:

۱- انتخاب مجموعه تصاویر مناسب از هدف: برای اینکه یک مجموعه داده‌ی تقریباً کاملی داشته

باشیم، تعدادی تصویر از هدف مورد نظر در حالت‌های مختلف را انتخاب می‌کنیم. به طوری

که تقریباً از تمام زوایا از هدف عکسی داشته باشیم.

۲- استخراج ویژگی: ویژگی‌های بکار رفته در پایان‌نامه شامل ویژگی رنگ و ویژگی بافت هستند.

از سه ویژگی رنگ (R,G,B) و ویژگی‌های میانگین و الگوی باینری دودویی به عنوان

ویژگی‌های بافت استفاده می‌کنیم. بنابراین تعداد ستون‌های بردار ویژگی برای آموزش

ماشین بردار پشتیبان پنج است. به دلیل اینکه ممکن است پس‌زمینه‌های موجود در تصاویر

انتخابی تمامی پس‌زمینه‌های موجود را دربر نگیرد، لازم است تا تعدادی داده‌ی اضافی به

عنوان پس‌زمینه به داده‌های آموزشی اضافه کنیم. این داده‌های اضافی باید در بازه‌ای قرار

گیرند که داده‌های استخراجی از تصاویر در این بازه قرار ندارند.

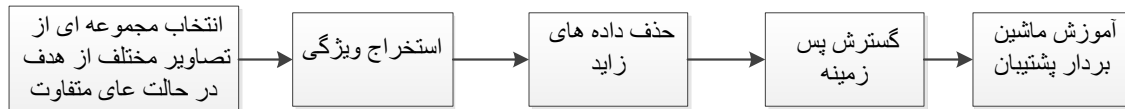
۳- کاهش تعداد داده‌ها: به دلیل اینکه مجموعه داده‌ی استخراجی بسیار بزرگ است و می‌دانیم

که زمان لازم برای آموزش ماشین بردار پشتیبان تابعی درجه سه از تعداد داده‌هاست.

بنابراین لازم است تا جایی که به مسئله کلاس‌بندی خدشه‌ای وارد نشود داده‌ها را کاهش

دهیم.

بعد از بدست آوردن داده‌های موثر و همچنین پارامترهای بهینه برای ماشین بردار پشتیبان، نوبت به آموزش داده‌ها می‌رسد. در شکل (۴-۱) مراحل بالا به صورت گام به گام توضیح داده شده است.

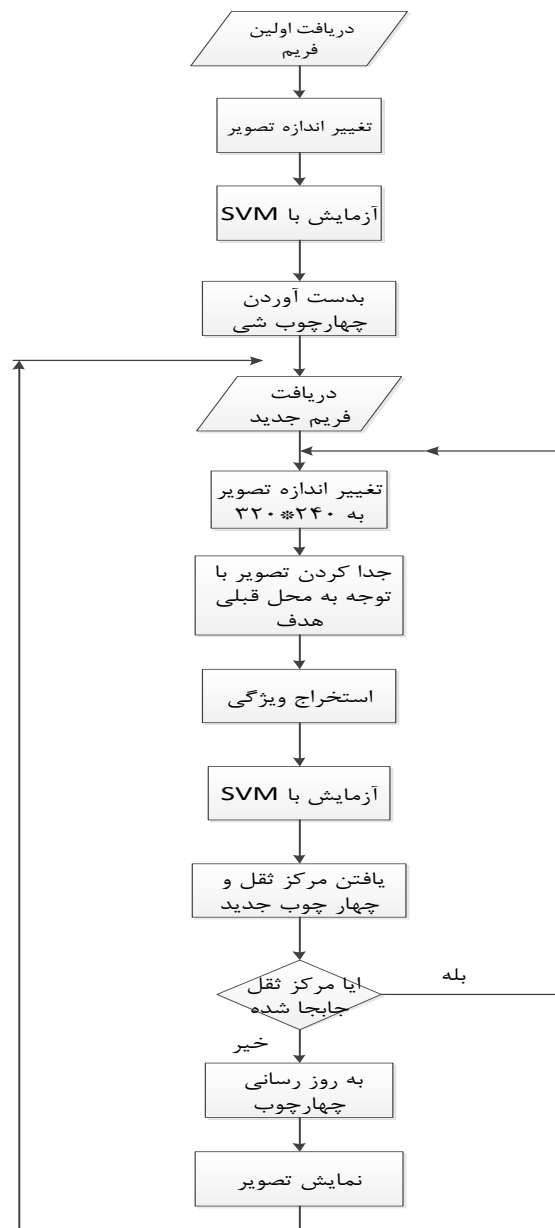


شکل (۴-۱)-مراحل مربوط به آموزش ماشین بردار پشتیبان

۴-۲- ردگیری هدف متحرک:

آشکارسازی هدف: اولین مرحله برای ردگیری هدف، آشکارسازی هدف مورد نظر است. برای این منظور بعد از دریافت اولین فریم و استخراج ویژگی‌های آن، ویژگی‌ها را به ماشین بردار پشتیبانی که به صورت آفلاین آموزش داده‌ایم، به عنوان ورودی آزمایش می‌دهیم. بعد از کلاس‌بندی فریم و برچسب‌گذاری پیکسل‌ها برای اینکه در فریم بعدی نیاز به آزمایش تمامی عکس نباشد و همچنین زمانی برای استخراج داده‌های زاید هدر نرود، لازم است فقط قسمتی از تصویر که هدف مورد نظر در آنجاست جستجو را انجام دهیم. برای نمایش هدف از روش شکل‌های هندسی اولیه استفاده می‌کنیم و با یک مستطیل هدف را محاط می‌کنیم. بعد از انتخاب محدوده هدف و تعیین مستطیل برای نمایش آن باید این مستطیل در هر فریم همراه هدف حرکت کند برای این کار بعد از ورود فریم جدید و تعیین محدوده هدف براساس چهارچوب (مستطیلی که توسط آن هدف را نمایش می‌دهیم) قبلی آن را با ماشین بردار پشتیبان آزمایش می‌کنیم و مرکز ثقل تصویر باینری را بدست می‌آوریم. اگر مکان مرکز ثقل در فریم فعلی نسبت به فریم قبلی تغییر داشت مرکز چهارچوب را به مرکز ثقل هدف انتقال می‌دهیم. این کار را آنقدر انجام می‌دهیم تا مرکز هدف و مرکز چهارچوب بر هم منطبق شوند. چون اندازه هدف در فریم‌های مختلف متفاوت است برای اینکه این تغییرات را نیز لحاظ کرده باشیم اگر نسبت تعداد پیکسل‌های مربوط به هدف به کل پیکسل‌های چهارچوب از یک مقداری کمتر یا بیشتر شد چهارچوب را بزرگ یا کوچک می‌کنیم.

در فلو چارت شکل (۲-۴) مراحل بالا را مشاهده می کنید.



شکل (۲-۴) - فلوچارت مربوط به مرحله تشخیص و ردگیری هدف

۴-۳-ویژگی‌های بکار رفته در پایان‌نامه:

۴-۳-۱-ویژگی رنگ:

ویژگی رنگ بکار رفته، ویژگی R,G,B است. که در فصل قبل توضیح داده شد.

۴-۳-۲-ویژگی بافت:

اگر تنها از هیستوگرام رنگ در جابجایی میانگین استفاده کنیم مشکلاتی از جمله

الف-اطلاعات فضایی هدف از بین می‌رود.

ب- وقتی هدف هم‌رنگ پس‌زمینه است هیستوگرام رنگ قادر به تشخیص هدف نخواهد بود

بوجود خواهد آمد. برای تشخیص بهتر هدف ویژگی‌های لبه را با ویژگی رنگ ترکیب می‌کنیم.

الگوهای بافت که ساختار فضایی شیء را منعکس می‌کنند، ویژگی‌های موثری برای تشخیص هدف هستند. از آنجایی که ویژگی‌های بافت اطلاعاتی را استخراج می‌کنند که ویژگی رنگ نمی‌تواند آن را استخراج کند بنابراین استفاده‌ی توأم از ویژگی رنگ و بافت از استفاده‌ی به تنهایی از ویژگی رنگ قابل اطمینان‌تر است. هر چند چگونگی استفاده‌ی موثر از ترکیب این ویژگی‌ها هنوز یک مسئله حل نشده است. با این وجود روش‌های زیادی نظیر، ماتریس همزمانی خاکستری و فیلتر گابور پیشنهاد شده‌اند که اینها دارای محاسبات سنگین می‌باشند و نمی‌توانند همزمان با هیستوگرام رنگ مورد استفاده قرار گیرند. در این پایان‌نامه از دو ویژگی بافت استفاده شده است که در زیر مختصری در مورد آن بحث می‌کنیم.

۴-۳-۲-۱- ویژگی بافت میانگین:

یکی از ویژگی‌های بافت، ویژگی میانگین است [۳۱]. برای یافتن این ویژگی ابتدا یک ماسک با ابعاد $n*n$ (که n معمولاً یک عدد فرد است) تعیین می‌کنیم. این ماسک را بر روی تصویر منطبق می‌کنیم و پیکسل معادل مرکز ماسک در تصویر را از فرمول زیر محاسبه می‌کنیم. سپس ماسک را بر روی تصویر حرکت می‌دهیم و سایر مقادیر را نیز به همین منوال محاسبه می‌کنیم.

$$Average = \sum_{i=1}^{n*n} w_i I_i \quad (۱-۴)$$

W_i : وزن اختصاص داده شده توسط ماسک

I_i : مقدار پیکسل i ام

در شکل (۴-۳) یک ماسک با ابعاد $3*3$ نشان داده شده است.

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

شکل (۴-۳) یک ماسک $3*3$ [۳۲]

در پایان‌نامه از یک ماسک $3*3$ با وزن‌های یکسان استفاده شده است. که در زیر فرمول محاسبه و شکل ماسک را ملاحظه می‌کنید.

$$Average = \frac{1}{9} \sum_{i=1}^9 I_i \quad (۲-۴)$$

$$\frac{1}{9} \times$$

1	1	1
1	1	1
1	1	1

شکل (۴-۴) ماسک استفاده شده در پایان نامه [۳۲]

۴-۳-۲- الگوی باینری محلی:

الگوی باینری محلی یک روش موثر برای توصیف ویژگی‌های بافت است. این روش مزایایی مانند، سرعت محاسبات بالا و تغییر ناپذیری با چرخش را دارد. و در زمینه‌هایی نظیر آنالیز بافت، بازیابی عکس، تشخیص چهره، قطعه‌بندی و غیره مورد استفاده قرار می‌گیرد [۲۲].

در الگوی باینری محلی برچسب گذاری به این صورت است که ابتدا یک همسایگی به شعاع R و تعداد P در نظر می‌گیریم هر کدام از پیکسل‌ها از پیکسل مرکزی بزرگتر بود آنرا یک و هر کدام کوچکتر بود آن را صفر می‌گیریم سپس با توجه به این مقادیر یک عدد باینری به پیکسل مرکزی نسبت می‌دهیم. در زیر روابط مربوط به این روش را می‌آوریم.

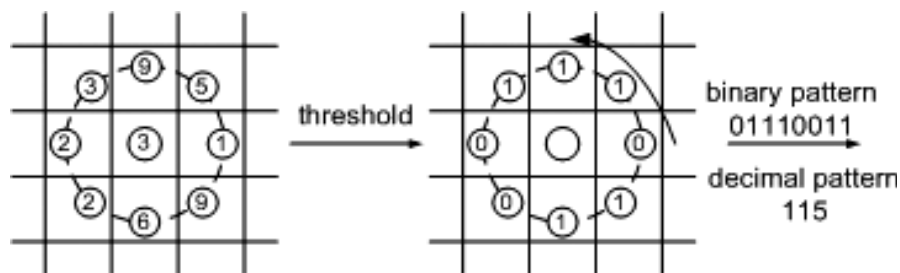
$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (۳-۴)$$

g_p و g_c : به ترتیب مقادیر مربوط به پیکسل مرکزی و پیکسل همسایه

(x_c, y_c) : مختصات پیکسل مرکزی

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (۴-۴)$$

شکل (۴-۵) یک مثال از LBP8,1 را نشان می‌دهد.



شکل (۵-۴) - یک مثال از $LBP_{8,1}$ [۲۲]

مدل بافتی که از روابط بالا نتیجه گرفته می‌شود فقط در مقیاس خاکستری تغییرناپذیر است. مدل بافت LBP که در برابر چرخش نیز تغییرناپذیر باشد از روابط زیر بدست می‌آید.

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P + 1 & \text{otherwise} \end{cases} \quad (۵-۴)$$

$$U(LBP_{P,R}) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (۶-۴)$$

۴-۴- جابجایی میانگین:

در این روش جابجایی شی با جابجایی مرکز ثقل چهارچوب باینری آن مدل می‌شود. روش کار به این صورت است که بعد از بدست آوردن تصویر باینری شی در چهارچوب، مرکز این شی را بدست می‌آوریم مرکز چهار چوب را منطبق با مرکز شی قرار می‌دهیم. دوباره توسط مختصات جدید چهارچوب هدف را از فریم جدا کرده و آنرا به شبکه می‌دهیم و دوباره مرکز شی را محاسبه می‌کنیم. این کار را آنقدر انجام می‌دهیم تا دیگر مرکز چهارچوب تغییر نکند. به عبارت دیگر شی بطور کامل در چهارچوب قرار گیرد. در هر تکرار مرکز ثقل شی داخل چهارچوب از معادله زیر محاسبه می‌شود.

$$X_{new} = \frac{\sum_{i=1}^N (I_i \cdot X_i)}{\sum_{i=1}^N I_i} \quad (۷-۴)$$

$$Y_{new} = \frac{\sum_{i=1}^N (I_i \cdot Y_i)}{\sum_{i=1}^N I_i} \quad (۸-۴)$$

که X_i, Y_i بیانگر محل پیکسل در تصویر است. و I_i مقدار پیکسل است به صورت که اگر جزء شی باشد مقدار آن یک و اگر جزء پس‌زمینه باشد مقدار صفر است. N تعداد پیکسل‌های درون چهارچوب است. در این روش لازم است که در دو فریم متوالی چهارچوب‌های شی با هم همپوشانی داشته باشند، که در حالت طبیعی این اتفاق می‌افتد. بعلاوه استفاده از شبکه svm اگر قسمتی از شی تحت همپوشانی قرار گیرد پیکسل‌های باقیمانده هنوز قابل تشخیص هستند. بنابراین این روش تحت همپوشانی دچار اختلال نخواهد شد. ولی در مواقعی که همپوشانی کامل داریم برای دنبال کردن شی در هنگام پوشیدگی احتیاج به روش‌هایی داریم که با وجود پوشیدگی شی را در فریم‌های بعدی پیش‌بینی کند. از جمله روش‌های که می‌توان از آن به عنوان پیش‌بینی‌کننده استفاده کرد، فیلتر کالمن یا فیلتر ذره‌ای می‌باشند.

۴-۵- روش خوشه‌بندی k_mean

این روش علی‌رغم سادگی آن یک روش پایه برای بسیاری از روش‌های خوشه‌بندی دیگر (مانند خوشه‌بندی فازی) محسوب می‌شود. این روش روشی انحصاری و مسطح محسوب می‌شود. برای این الگوریتم شکل‌های مختلفی بیان شده است. ولی همه آنها دارای روالی تکراری هستند که برای تعدادی ثابت از خوشه‌ها سعی در تخمین موارد زیر دارند:

۱- بدست آوردن نقاطی به عنوان مراکز خوشه‌ها این نقاط در واقع همان میانگین نقاط متعلق

به هر خوشه هستند.

۲- نسبت دادن هر نمونه داده به یک خوشه که آن داده کمترین فاصله تا مرکز آن خوشه را

دارا باشد.

در نوع ساده‌ای از این روش ابتدا به تعداد خوشه‌های مورد نیاز نقاطی به صورت تصادفی انتخاب می‌شود. سپس در داده‌ها با توجه با میزان نزدیکی (شباهت) به یکی از این خوشه‌ها نسبت داده

می‌شوند و بدین ترتیب خوشه‌های جدیدی حاصل می‌شود. با تکرار همین روال می‌توان در هر تکرار با میانگین‌گیری از داده‌ها مراکز جدیدی برای آنها محاسبه کرد و مجدداً داده‌ها را به خوشه‌های جدید نسبت داد. این روند تا زمانی ادامه پیدا می‌کند که دیگر تغییری در داده‌ها حاصل نشود.

در زیر روال مرحله به مرحله الگوریتم بیان می‌شود. [۲۸]

فرض: m : تعداد داده‌ها و l : تعداد خوشه‌ها

مرحله ۱- انتخاب مقدار l به طوری که $l < m$

مرحله ۲- انتخاب l تا از داده‌ها به عنوان مراکز خوشه‌ها $c_j = x_j \quad j = 1, 2, \dots, l$;

مرحله ۳- تخصیص $x_i \quad i=1, l+1, \dots, m$ به یکی از کلاس‌ها که کمترین فاصله تا مرکز آن خوشه را دارد. یعنی x_j متعلق به خوشه j^* است اگر

$$\|x_i - c_{j^*}\| = \min_{j^*} \|x_i - c_j\|, \quad 1 \leq j \leq l \quad (9-4)$$

مرحله ۴- محاسبه مجدد مراکز خوشه‌ها با استفاده از میانگین داده‌های مربوط به خوشه

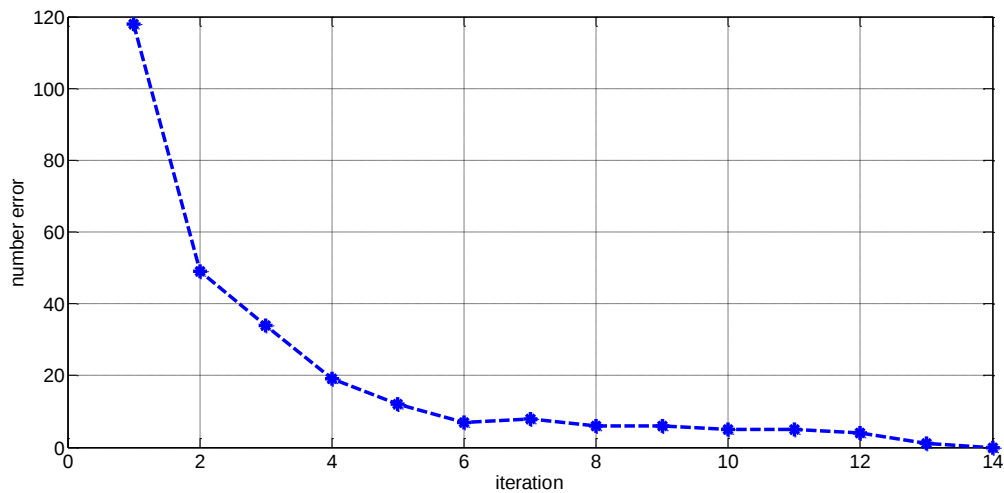
$$c_j = \frac{1}{m_j} \sum_{i \in c_j} x_i; \quad 1 \leq j \leq l \quad (10-4)$$

که m_j : تعداد داده‌های مربوط به خوشه j ام

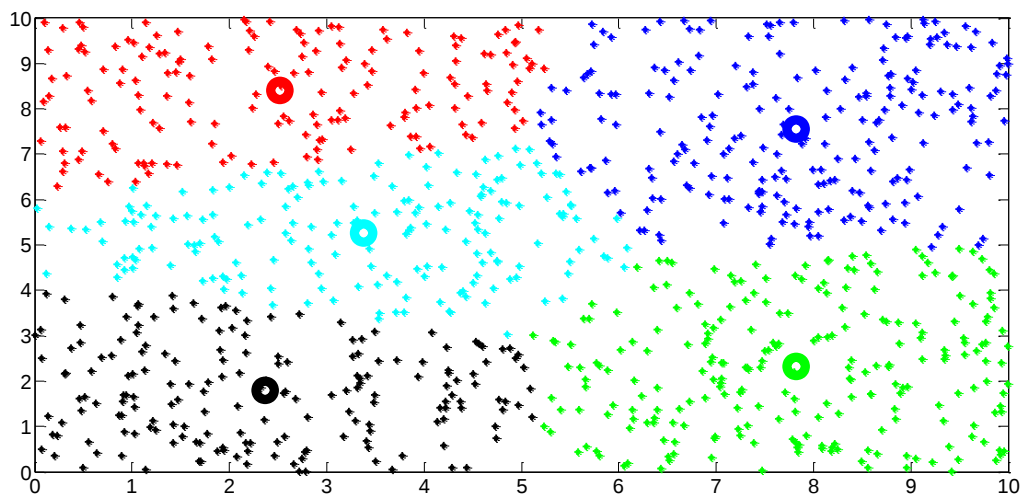
مرحله ۵- تعیین کنید $x_i \quad i=1, 2, \dots, m$ بر اساس نزدیکترین فاصله تا مراکز خوشه‌ها، مربوط به کدام خوشه است.

مرحله ۶- اگر حداقل یکی از داده‌ها اشتباه کلاس‌بندی شده بود به مرحله ۴ برگرد در غیر این صورت پایان فرایند خوشه‌بندی.

در شکل‌های زیر چگونگی به صفر رسیدن خطا و رسیدن به خوشه‌بندی درست نمایش داده شده است. برای نمایش راحت‌تر از داده‌های دوبعدی و تعداد مراکز کلاس را ۵ در نظر می‌گیریم. تعداد داده‌ها ۱۰۰۰ است که بصورت رندوم در بازه‌ی [۰ ۱۰] انتخاب می‌شوند.



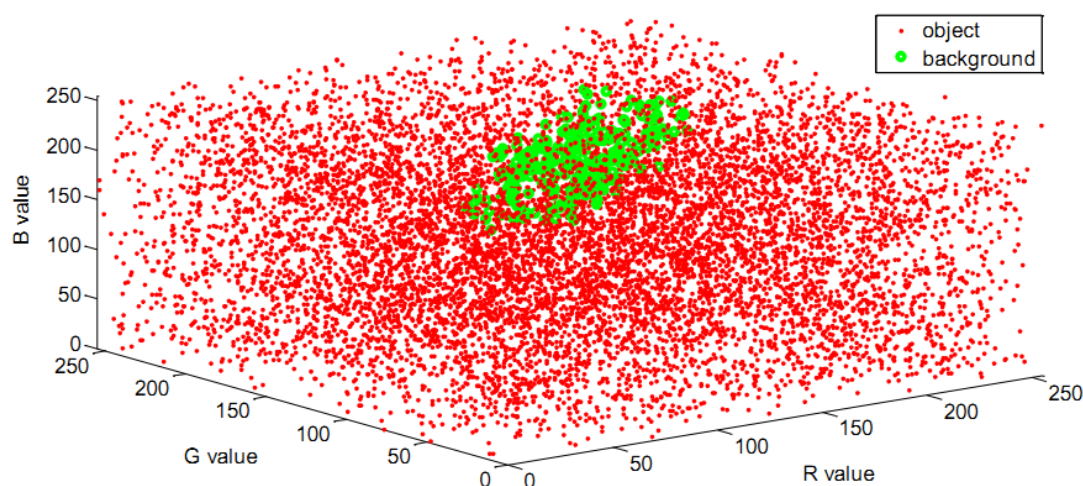
شکل (۴-۶) - روند به صفر رسیدن تعداد داده‌های اشتباه خوشه‌بندی شده



شکل (۴-۷) - خوشه‌بندی داده‌ها در پنج کلاس

۴-۵- گسترش پس زمینه

به دلیل اینکه ممکن است تصاویر انتخابی تمامی پس زمینه‌های لازم را در برنگیرد، لازم به توسعه پس زمینه داریم. همچنین اگر پس زمینه را گسترش ندهیم با توجه به اینکه داده‌های مربوط به هدف در فضای ویژگی (برای نمایش فقط از فضای رنگ استفاده می‌کنیم) طوری است که اگر یک شیء دیگر دارای رنگی غیر از رنگ‌های بکار رفته شده در تصاویر انتخابی باشد به احتمال زیاد هدف در نظر گرفته می‌شود (البته این موضوع برای فضای ویژگی رنگ بررسی شده و معلوم شد که داده‌های خارج از مجموعه داده، به عنوان هدف در نظر گرفته می‌شود) [۳۳]. این موضوع در شکل (۴-۸) نمایش داده شده است. بنابراین لازم است که پس زمینه را گسترش دهیم.



شکل (۴-۸) آزمایش تعدادی داده در فضای ویژگی قبل از گسترش پس زمینه

روش کار توسعه پس زمینه:

در این روش به تعدادی دلخواه داده در فضای خارج از فضای مجموعه داده و در فضای رنگ (یک فضای سه بعدی که هر بعد آن در بازه‌ی [۰ ۱] قرار دارد) در نظر می‌گیریم. برای این کار به روش زیر عمل می‌کنیم.

ابتدا برای اینکه داده‌های انتخابی با داده‌هایی که به عنوان ویژگی‌های هدف می‌باشند تداخلی نداشته باشد، داده‌های مربوط به هدف را با استفاده از روش K-mean خوشه‌بندی می‌کنیم. بعد فاصله دورترین داده مربوط به هر خوشه را بدست می‌آوریم. حال یک عدد دلخواه انتخاب می‌کنیم، فاصله این عدد را تا نزدیکترین مرکز خوشه مربوط به هدف بدست می‌آوریم. اگر این فاصله بیشتر از فاصله دورترین نقطه‌ی آن خوشه بود آن را به عنوان داده انتخابی برای گسترش پس‌زمینه در نظر می‌گیریم. در شکل زیر داده‌های آموزشی برای ویژگی رنگ نشان داده شده است.

مراحل گسترش پس‌زمینه در زیر مرحله به مرحله توضیح داده شده است.

مرحله ۱- انتخاب K, N به ترتیب تعداد داده‌های مورد نظر برای گسترش پس‌زمینه و تعداد خوشه‌های مربوط به داده‌های هدف

مرحله ۲- استفاده از روش K_mean و تعیین مراکز هر خوشه

مرحله ۳- تعیین دورترین نقطه به مرکز هر خوشه و فاصله این نقطه تا مرکز خوشه

مرحله ۴- انتخاب یک بردار دلخواه در ابعاد $1 \times m$ در فضای ویژگی که m در این پایان‌نامه برابر ۵ است.

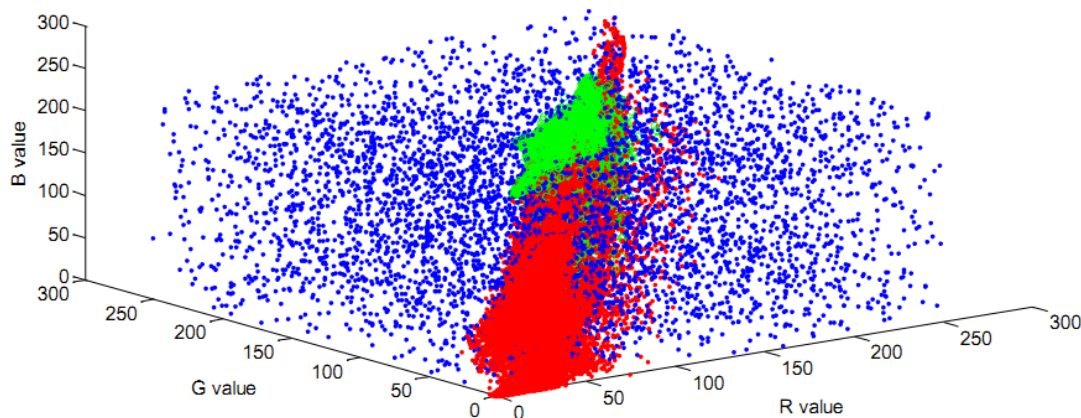
مرحله ۵- تعیین نزدیکترین مرکز خوشه به بردار مرحله ۴

مرحله ۶- اگر فاصله بردار تا مرکز خوشه بیشتر از دورترین نقطه مربوط به خوشه بود، این بردار را به

عنوان ویژگی گسترش پس‌زمینه در نظر بگیر. و $N=N-1$;

مرحله ۷- اگر $N \geq 1$ انگاه به مرحله ۴ برو

در شکل (۴-۹) داده‌های استخراج شده از عکس‌های منتخب به همراه داده‌هایی که از گسترش پس‌زمینه بدست آمده نشان داده شده است.



شکل (۴-۹) داده‌های استخراج شده از عکس‌های انتخابی و داده‌های بدست آمده از گسترش پس‌زمینه

۴-۶- بروز رسانی تغییرات در اندازه

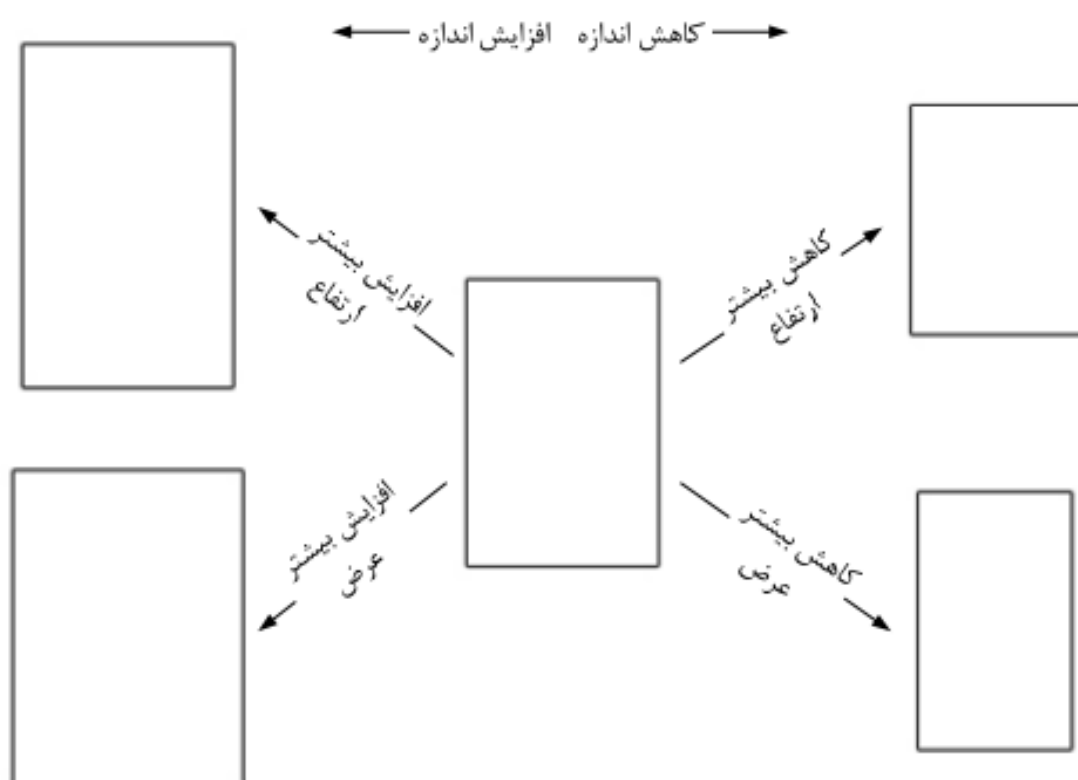
در حالت کلی در طول ردیابی اندازه‌ی ثابتی نخواهد داشت. برای جبران این حالت انجام بروز رسانی اندازه شیء اجتناب ناپذیر است.

همانطور که گفته شد در طول ردیابی شیء اندازه ثابتی نخواهد داشت. برای بروز رسانی اندازه شیء ابتدا در فریم اول نسبت مجموع پیکسل‌های شیء نسبت به مساحت چهارچوب شیء را طبق رابطه زیر محاسبه می‌کنیم.

$$\text{نسبت مساحت} = \frac{\sum_{i=1}^N I_i}{N} \quad (۴-۱۱)$$

که در آن I_i مقدار پیکسل در تصویر باینری می‌باشد که می‌تواند ۰ یا ۱ باشد. N تعداد پیکسل‌های درون چهارچوب است. از این معیار برای بدست آوردن چهارچوب مناسب برای شیء استفاده می‌شود. برای مقاومت در برابر اشتباهات جزئی و نویز تنها هنگامی که تفاوت مقادیر نسبت مساحت جدید و اصلی بیش از ۷.۵٪ باشد بروز رسانی اندازه انجام خواهد گرفت. اگر نسبت مساحت محاسبه شده کمتر از ۹۲.۵٪ حالت قبل گردد به این معناست که شیء کوچکتر شده و بنابراین پنجره‌ی شیء باید کوچک شود. در پروسه کوچک کردن پنجره برای آن که بتوان تغییرات در نسبت ابعاد شیء را نیز

جبران کرد پنجره‌ی شیء در دو وضعیت کاهش می‌یابد. یک بار عرض پنجره به میزان $\frac{1}{40}$ عرض قبلی و طول پنجره به میزان $\frac{1}{20}$ طول قبلی کاهش می‌یابد (طول کاهش بیشتری می‌یابد) و بار دیگر عرض پنجره به میزان $\frac{1}{20}$ عرض قبلی و طول پنجره به میزان $\frac{1}{40}$ طول قبلی کاهش می‌یابد (عرض کاهش بیشتری می‌یابد) و نسبت مساحت برای هر دو حالت محاسبه می‌شود. وضعیتی که نسبت مساحت نزدیکتری به نسبت مساحت اصلی داشت به عنوان نسبت مساحت جدید انتخاب می‌گردد. باید توجه داشت که در هر دو حالت میزان کاهش برابر است و مساحت در وضعیت جدید طبق رابطه (۴-۱۲) برابر با ۹۲.۵٪ حالت قبل خواهد بود که جبران کننده‌ی کاهش اندازه‌ی شیء است. پروسه مشابهی برای جبران اندازه‌ی چهارچوب شیء در هنگامی که شیء بزرگتر می‌شود انجام می‌گیرد. که طبق رابطه (۴-۱۳) ثابت می‌گردد که جبران کننده‌ی میزان اندازه شیء است. در حالت کلی ۴ حالت وجود دارد که در شکل (۴-۱۰) نمایش داده شده است.



شکل (۴-۱۰) حالاتی که برای به روز رسانی اندازه پیش می‌آید.

$$\text{مساحت کاهش یافته} = \left(w - \frac{1}{40}w\right) \left(h - \frac{1}{20}h\right) = \left(w - \frac{1}{20}w\right) \left(h - \frac{1}{40}h\right) = 0.925(w * h) \quad (۱۲-۴)$$

$$\text{مساحت افزایش یافته} = \left(w + \frac{1}{40}w\right) \left(h + \frac{1}{20}h\right) = \left(w + \frac{1}{20}w\right) \left(h + \frac{1}{40}h\right) = 1.075(w * h) \quad (۱۳-۴)$$

البته روش بالا برای حالتی است که تغییرات در حالت افزایش اندازه شیء در دو فریم متوالی به اندازه-ای نباشد که مرزهای شیء از چهارچوب خارج شود. که در عمل معمولاً چنین اتفاقی نمی‌افتد.

فصل پنجم

نتایج شبیه‌سازی، نتیجه‌گیری و پیشنهادات

مقدمه

در این قسمت به بررسی عملکرد الگوریتم پیشنهادی می‌پردازیم. در ابتدا خروجی هر مرحله از الگوریتم را نمایش می‌دهیم و سپس آن را با دو مقاله در همین زمینه مقایسه می‌کنیم. برنامه این مقالات از سایت نویسندگان آن دانلود شده، بنابراین مقایسه انجام شده قابل قبول است [۳۴]. این مقایسه در زمینه شباهت مسیر پیموده شده توسط الگوریتم‌ها با میسر واقعی و مدت زمان لازم جهت اجرای الگوریتم می‌باشد. تفاوت این کار با کار نویسندگان مقاله‌هایی که قرار است با آن مقایسه انجام گیرد در این است که آن‌ها آموزش آنلاین دارند.

۵-۱- بدست آوردن مسیر واقعی:

برای بدست آوردن مسیر واقعی قطعه‌ی مورد نظر را انتخاب می‌کنیم و با روشی دلخواه برای هر فریم آن تصویر باینری مربوط را بدست می‌آوریم سپس قسمت‌هایی از عکس که اشتباه تشخیص داده شده را درست می‌کنیم. در این زمان مرکز ثقل تصویر باینری را توسط فرمول زیر محاسبه می‌کنیم.

$$x_{center} = \frac{\sum X_i \cdot I}{\sum I} \quad (۱-۵)$$

$$y_{center} = \frac{\sum Y_i \cdot I}{\sum I} \quad (۲-۵)$$

I: تصویر باینری مربوط به فریم

Y_i : یک ماتریس که درایه‌های هرستونش برابر با شماره آن ستون است

X_i : یک ماتریس که درایه‌های هرسطرش برابر با شماره آن ستون است

در شکل زیر یک فریم مورد آزمایش به همراه تصویر باینری و مرکز ثقل آن نشان داده شده است.



a



b

شکل (۵-۱). تصویر به همراه شکل باینری آن و مرکز ثقل آن

به همین طریق برای سایر فریم‌ها مرکز ثقل هدف را تعیین می‌کنیم.

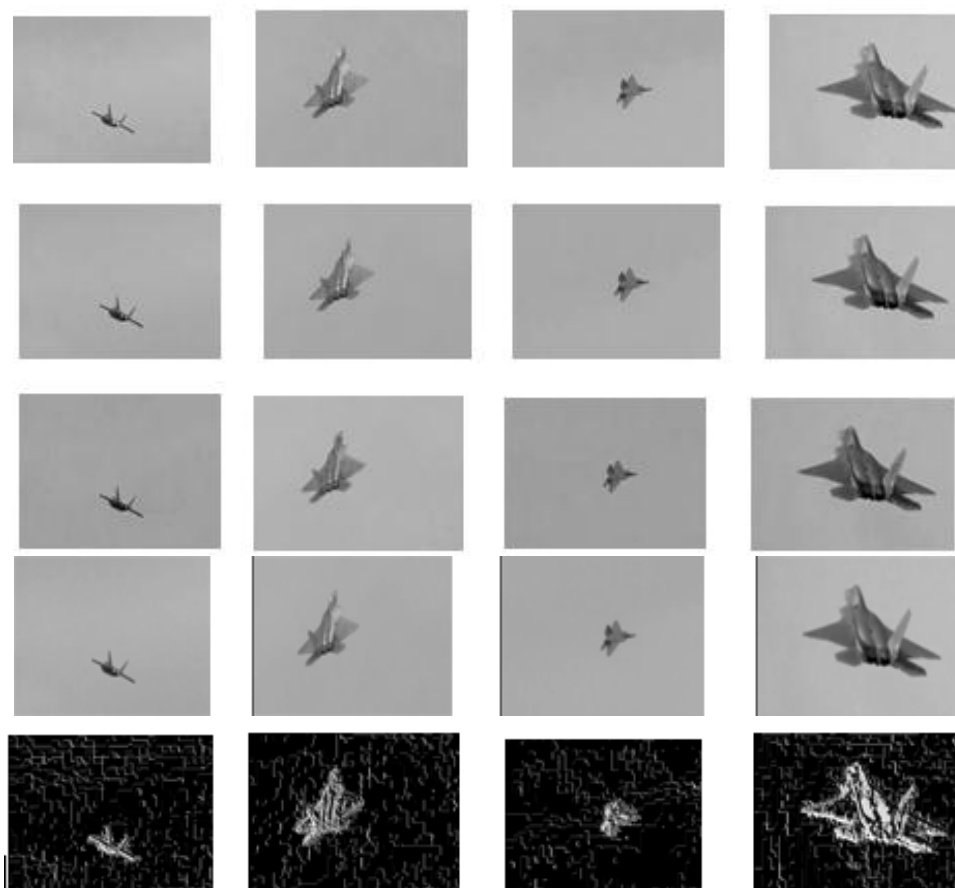
حال نوبت به نمایش خروجی تصاویر بعد از هر مرحله از الگوریتم است. اولین گام برای ردگیری هدف یافتن یک مجموعه داده برای آموزش ماشین بردار پشتیبان و جدا کردن تصاویر از زمینه است و سپس یافتن تصویر دودویی مربوط به تصاویر است. در شکل زیر تعدادی از تصاویری که برای آموزش بکار رفته‌اند نمایش داده شده است.



شکل (۵-۲). سطر اول تصاویر منتخب برای آموزش. سطر دوم تصاویر متناظر با سطر اول که هدف بصورت دستی جدا شده. تصاویر باینری مربوط که دارای برچسب یک و صفر است.

۵-۲- استخراج ویژگی

ویژگی‌های بکار رفته در پایان‌نامه شامل ویژگی رنگ (R,G,B) و ویژگی میانگین و الگوی دودویی محلی است که در زیر ویژگی استخراج شده از هر نوع را مشاهده می‌کنید.



شکل (۵-۳). ویژگی‌های استخراج شده از تصاویر شکل (۵-۲)، ویژگی‌های رنگ R ، G ، B ، الگوی دودویی محلی و میانگین به ترتیب در سطرها یک تا پنج

۵-۳- ردگیری هدف

همانطور که گفته شد برای اینکه چهارچوب با هدف حرکت کند لازم است تا مرکز مستطیل با مرکز ثقل هدف مطبق شود. در شکل زیر چگونگی منطبق شدن نمایش داده شده است. در شکل (۵-۴) خروجی ماشین بردار پشتیبان برای اولین فریم و قسمتی که به عنوان هدف تشخیص داده شده است، نمایش داده شده است.



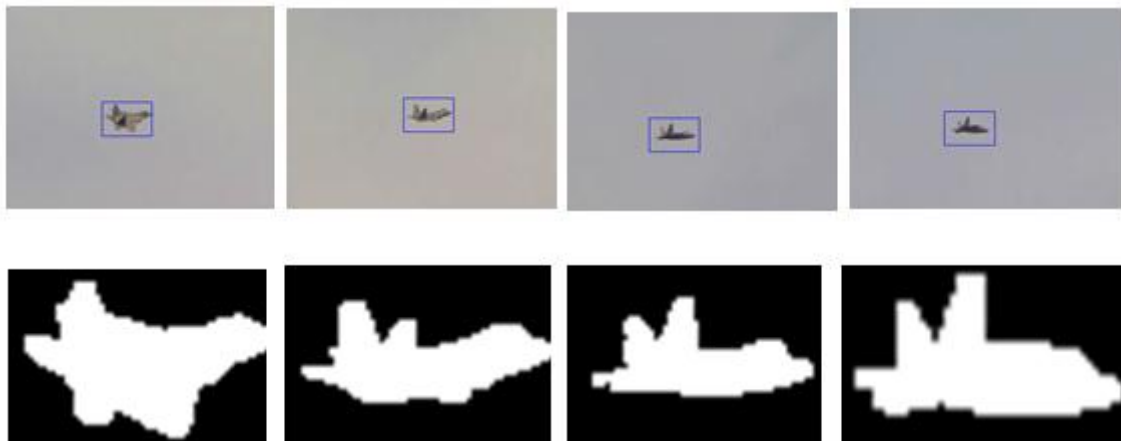
شکل (۴-۵). به ترتیب از چپ به راست، اولین فریم از فیلم، تصویر باینری معادل آن که خروجی ماشین بردار پشتیبان است و قسمتی از تصویر که به عنوان هدف تشخیص داده شد.



شکل (۵-۵). چهارچوب کم رنگ مربوط به فریم قبل و چهارچوب تیره مربوط به فریم فعلی

همانطور که مشاهده می شود بعد از یک با تکرار مرکز چهارچوب و مرکز ثقل بر هم منطبق شده اند.

در انتها نیز چند فریم به همراه قسمت جدا شده توسط چهارچوب مربوط به آن آمده است.

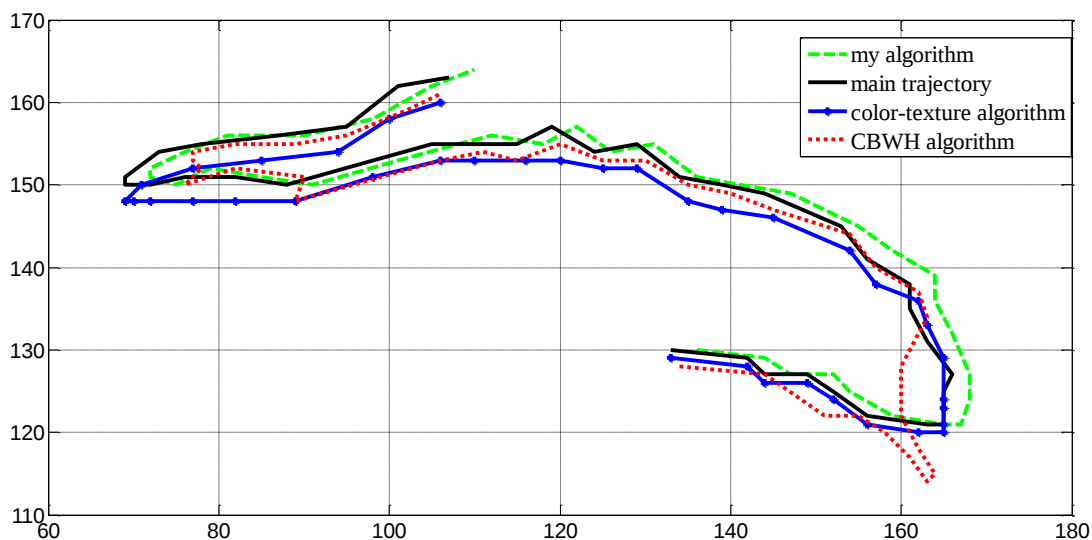


شکل (۵-۶). فریم‌های ۲۰، ۵۰، ۱۰۰ و ۱۵۰ به همراه چهارچوب باینری مربوط به آن در انتها روش پیشنهادی را با دو کار دیگر مقایسه کنیم. در مقاله اول [۲۲] از هیستوگرام رنگ و بافت^۱ برای نمایش هدف تواما استفاده کرده است. ویژگی بافت بکار رفته در این مقاله استفاده از الگوی دودویی محلی است. و روش جابجایی میانگین برای ردگیری استفاده کرده است. و در [۳۵] از الگوریتم هیستوگرام پس‌زمینه وزن‌دار شده تصحیح شده^۲ برای بازنمایی هدف و الگوریتم جابجایی میانگین برای ردگیری استفاده شده است. همچنین تنها از ویژگی رنگ برای شناسایی هدف استفاده کرده است.

در شکل (۵-۶) اختلاف مسیر حرکت واقعی هدف را با سایر الگوریتم‌ها مشاهده می‌کنیم. و در جدول (۵-۱) تفاوت در سرعت الگوریتم‌ها آورده شده است.

^۱ Color_texture Histogram

^۲ Corrected Background-Weighted Histogram(CBWH)



شکل (۷-۵) اختلاف مسیر طی شده‌ی هدف در الگوریتم‌های متفاوت با مسیر اصلی

جدول (۱-۵). مقایسه زمان اجرای الگوریتم پیشنهادی با الگوریتم‌های معرفی شده در بالا

<i>algorithm</i>	<i>My algorithm</i>	<i>CBWH Algorithm</i>	<i>Color-texture</i>
<i>Frame/sec</i>	<i>2.67</i>	<i>3.94</i>	<i>2.56</i>

همانطور که از شکل (۷-۵) و جدول (۱-۵) مشخص است الگوریتم پیشنهادی در سرعت و دقت بر الگوریتم ارائه شده در [۳۵] برتری دارد. و همچنین از لحاظ دقت از الگوریتم ارائه شده در [۲۲] بهتر ولی دارای سرعت پایین‌تری نسبت به این الگوریتم است.

۵-۴- نتیجه گیری:

در این پایان نامه برای بهبود مدل هدف، استفاده‌ی توأماً ویژگی رنگ و بافت را بکار برده‌ایم. و از جابجایی میانگین برای ردگیری استفاده کرده‌ایم. و همچنین ماشین بردار پشتیبان را به صورت افلاین آموزش داده و برای جداسازی هدف از پس‌زمینه مورد استفاده قرار می‌دهیم. نتایج آزمایش نشان می‌دهد که بکارگیری ویژگی‌های رنگ و بافت با هم بهتر است از بکارگیری ویژگی رنگ به تنهایی است. هرچند به دلیل اینکه لازم است تا ویژگی‌های بافت را نیز استخراج کنیم زمان پردازش طولانی تر می‌شود. ولی در عوض دقت پردازش بالاتر می‌رود.

۵-۵- پیشنهادها:

اگر یک شیء دارای ویژگی‌های مشابه هدف از کنار آن عبور کند ممکن است الگوریتم دچار اشتباه شود و هدف را گم کند. برای جلوگیری از این موضوع می‌توان از روش‌های پیش‌بینی از جمله فیلتر کالمن استفاده کرد تا مکان آینده‌ی هدف را پیش‌بینی کرده و در مکان پیش‌بینی شده هدف را جستجو کنیم. و همچنین می‌توان به جای استفاده از ویژگی رنگ R,G,B از HSV استفاده کرد که دارای خاصیت یکنواختی ادراکی (متناظر با درک و تجربه انسان از رنگ) است. همچنین می‌توان برای بدست آوردن ضرایب بهینه ماشین بردار پشتیبان از الگوریتم‌های هوشمند مانند PSO، الگوریتم ژنتیک، الگوریتم کلونی مورچگان و غیره استفاده کرد.

منابع

- [1] A. Yilmaz, O. Javed, and M. Shah, (2006), "Object tracking: A survey", *ACM Computing Surveys*, Vol. 38, No. 4, pp. 1–45.
- [2] I. S. Kim, H. S. Choi, K. M. Yi, J. Y. Choi and S. G. Kong, (2010), "Intelligent Visual Surveillance - A Survey", *International Journal of Control Automation and Systems*, Vol. 8, No. 5, pp. 926–939.
- [3] K. Mikolajczyk and C. Schmid, (2003), "A Performance Evaluation of Local Descriptors", in: *Proc. IEEE Conference on Computer Vision And Pattern Recognition (CVPR)*, pp. 1615–1630.
- [4] B. D. Lucas and T. Kanade, (1981) , "An Iterative Image Registration Technique with An Application To Stereo Vision", in: *Proc. International Joint Conference on Artificial Intelligence*.
- [5] H. Grabner, C. Leistner, and H. Bischof, (2008), "Semi-supervised On-Line Boosting for Robust Tracking", in: *Proc. European Conference on Computer Vision (ECCV)*, Vol.1, pp. 234-247.
- [6] P. Fieguth and D. Terzopoulos, (1997), "Color-Based Tracking of Heads and Other Mobile Objects at Video Frame Rates", in: *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21–27.
- [7] D. Comaniciu, V. Ramesh and P. Meer, (2003), "Kernel-based object tracking", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, pp. 564–575.
- [8] D. Comaniciu, (2002), "Bayesian kernel tracking", in: *Proc. Annual Conference of the German Society for Pattern Recognition*, pp. 438–445.
- [9] R. Collins, Y. Liu and M. Leordeanu, (2005), "On-Line Selection of Discriminative Tracking Features", *IEEE Transactions on Pattern Analysis And Machine Intelligence*, Vol. 27, No. 10, pp.1631–1643.
- [10] S. Avidan, (2007), "Ensemble tracking", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 2, pp. 261–271.
- [11] C. Hua, H. Wu, Q. Chen and T. Wada, (2006), "K-means tracker: A General Algorithm for Tracking People", *Journal of Multimedia*, Vol.1, No. 4, pp. 46-53.

- [12] A. P. Kumar, T. N. Ricktesh, R.V. Babu, R. Hariharan, (2009), "Object Tracking Using Radial Basis Function Networks", National Aerospace Laboratories, No. Y-8-101.
- [13] S. Avidan, (2004), "Support Vector Tracking", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 8.
- [14] M. Bouzenada, M.C. Batouche, Z. Telli, (2007), "Neural network for object tracking", Information Technology Journal, Vol. 6, No. 4, pp. 526–533.
- [15] D. Huttenlocher, J. Noh and W. Rucklidge, (1993), "Tracking nonrigid objects in complex scenes" in: Proc. IEEE International Conference on Computer Vision (ICCV), pp.93–101.
- [16] F. Hausdorff, "Set Theory," Chelsea, New York, NY, 1962.
- [17] A. Yilmaz, X. Li and M. Shah, (2004), "Contour based object tracking with occlusion handling in video acquired using mobile cameras", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 11, pp. 1531–1536.
- [18] Mark S. Nixon, Alberto S. Aguado, (2002), "Feature Extraction and Image Processing", Newnes, 1st edition
- [19] Chung-Cheng Chiu, Min-Yu Ku, (2009)," Robust Background Subtraction Algorithm in Intelligence Traffic System", JMT, Vol. 28 No. 1 pp. 55-76
- [20] Ahmed Elgammal , David Harwood, Larry Davis , " Nonparametric Model for Background Subtraction", Computer Vision University of Maryland, College Park, MD 20742
- [21] R. Haralick, B. Shanmugam and I. Dinstein, (1973), "Textural features for image Classification", IEEE Transactions on Systems, Man, and Cybernetics, Vol. 33, No.3, pp. 610–622.
- [22] J. Ning, L. Zhang, D. Zhang and C. Wu, (2009), "Robust Object Tracking using Joint Color-Texture Histogram", International Journal of Pattern Recognition and Artificial Intelligence, Vol. 23, No. 7, pp. 1245–1263.
- [23] L. Nanni, S. Brahnam, A. Lumini, (2012), "Survey on LBP based texture descriptors for image classification", Expert Systems with Applications, Vol.39, No.3, pp.3634-3641.
- [24] J. Canny, (1986), "A computational approach to edge detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, pp. 679–698.
- [25] Duda R. O., Stork D. G., Hart P. E., (2001), "Pattern classification", John Wiley and Sons.

[26] جهانشاه کیودیان ، محمد رحمتی، محمد مهدی همایون پور، ۱۳۸۲، "استفاده از ماشین بردار پشتیبان در سه مساله شناسایی الگو"، اولین کنفرانس بین المللی فناوری اطلاعات و دانش.

- [27].Tristan Fletcher, (2009), "Support Vector Machine Explained", UCL, UK London
- [28] Madan M. Gupta, Liang Jin, and Noriyasu Homma, (2003), "*Static and Dynamic Neural Networks*", *John Wiley and Sons*, Canada.
- [29] Chih-Chung Chang and Chih-Jen Lin, (2011), " LIBSVM : A Library for Support Vector Machines", *TIST*, Volume 2 Issue 3, Article No. 27
- [30] Boyang LI, Qiangwei WANG and Jinglu HU, (2009), " A Fast SVM Training Method for Very Large Datasets", *IJCNN*, Atlanta, Georgia, USA,1784-1789
- [31] S.Selvarajah and S.R. Kodituwakku, (2011), " Analysis and Comparison of Texture Features for Content Based Image Retrieval", *IJLTC*, Volume 2, Issue 1, 108-113
- [32] Rafael C.Gonzalez, Richard E.Woods, (2002), "*Digital Image Processing*", Second Edition, Upper Saddle River, New Jersey.
- [33] Pradeep K. Deshmukh, Yogini Gholap, (2012), "Efficient Object Tracking Using K means and Radial Basis Function", *IJARCCCE*, Vol. 1, Issue 1, 27 – 31
- [34] <http://www4.comp.polyu.edu.hk/~cslzhang/papers.htm>
- [35]J. Ning, L. Zhang, D. Zhang and C. Wu, (2012), "Robust Mean Shift Tracking with Corrected Background-Weighted Histogram", *Computer Vision, IET*, Volume:6, Issue: 1.

Abstract

In this thesis, an algorithm for tracking a flying object is proposed based on using joint color_texture feature to represent a target. The Mean Shift algorithm is used for object tracking. In addition to the typical color feature, texture feature such as average feature and local binary pattern are taken into consideration for object identification. Moreover, a Support Vector Machine (SVM) is being used in order to distinguish between an object and its background. The task of learning SVM is accomplished using forty photos that objects and backgrounds separated manually, each of which from different angles. It is imaginable that some possible backgrounds were not included in training data during the tracking phase. In this case, to alleviate the problem in hand, background extension has been applied. As long as there are a great number of samples for training the SVM, the data which does not play a role in determining the hyper-plane has to be eliminated, and the training phase should be commenced afterwards. After receiving the first frame, one can test it with the SVM and express it in a framework subsequently, as well as using Mean Shift algorithm for tracking that frame in the next ones. The simulation results confirm the superiority of the proposed algorithm in both accuracy and speed over the traditional methods.

Keywords: Object Tracking, Support Vector Machine, Machine Vision, Local Binary Pattern, Mean Shift, Background Extension



Shahrood University of Technology
Department of Electrical and Robotic Engineering

Moving Target Tracking using Support Vector Machine(SVM)

By:

Reza mighani

Supervisor:

H.Tosian.Shandiz