

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد هوش مصنوعی

ردیابی زمان واقعی عابرین پیاده به صورت فشرده

نگارنده: سجاد جعفری

استاد راهنما:

دکتر حمید حسن پور

دکتر وحید ابوالقاسمی

شهریور ۹۷



فرم شماره (۳) صورتجلسه ایاب و ذهاب از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، این جلسه در جلسه دفاع از پایان نامه کارشناسی ارشد آقای سجاد جعفری دانشجو، دانشجو ۹۱۰۱۹۸۵ رشته مهندسی کامپیوتر، گزارش خوش منصفانه تحت عنوان «طراحی زبان واقعی نظریین پیاده به صورت فشرده» که در تاریخ ۱۳۹۳/۰۸/۱۸ با حضور هیأت محترم داوران در دانشگاه سمنان شامروید برگزار گردید و به شرح ذیل اجراء می‌گردد.

برود
 قبول با توجه به شرح زیر:

نوع تعلیم: نظری عملی

ردیف	حرفه جعفری	نام و نام خانوادگی	عضو هیأت داوران
	استاد	دکتر حمید حسن پور	۱- استاد دانشیار، پهل
	استاد	دکتر ویرانه نیکبخت	۲- استاد دانشیار، موم
			۳- استاد دانشیار
	استاد	دکتر ناصر فرهادی	۴- نماینده تحصیلات تکمیلی
	استاد	دکتر منصور حاج	۵- استاد دانشیار، پهل
	استاد	دکتر ناصر کرمانی	۶- استاد دانشیار، موم

به و نام خانوادگی: **سجاد جعفری**
 تاریخ و امضاء: **۱۳۹۳/۰۹/۰۱**

این فرم در صورتی که توسط داوران امضاء شود، جایگزین گزارش داوران می‌گردد. در صورتی که این فرم به هیچ وجه تکمیل نگردد، سمنان در تاریخ ۱۳۹۳/۰۹/۰۱ اعلام می‌گردد.

باسپاس از دو وجود مقدس:

آمان که ناتوان شد تا ما به توانایی برسیم...

موباشان سفید شد تا ما رو سفید شویم...

و عاشقانه سوختند تا کرم با بخش وجود ما و روشنگر راهمان باشند...

پدرانمان

مادرانمان

پروردگارا، نه میتوانم موباشان را که در راه عزت من سفید شد، سیاه کنم و نه برای دستهای پینه بسته شان که شمره تلاش برای افتخار من است، مرهمی دارم. پس توفیقم ده که هر لحظه شکر گزارشان باشم و ثانیه های عمرم را در عصای دست بودشان بگذرانم.

تشکر و قدردانی

سپاس و ستایش مخصوص خالق بی‌همتایی است که الطاف بی‌کرانش بجز زندگی ام را دربر گرفته و فضل بی‌قیاسش در مراحل دشوار زندگانی شامل عالم گشته که اگر عنایت بی‌منت او نبود، رساندن این بار به سر منزل مقصود غیر ممکن می‌نمود. اکنون که این پژوهش به زیور چاپ آراسته می‌گردد بر خود لازم می‌دانم تا مطابق سنت حسنه پاسکزاری، والاترین مراتب سپاس خویش را به محضر استادان فرزانه‌ای که در طول تحصیل، همواره افتخار شاگردی آنان را داشته‌ام و مرا م侖ون الطاف خالصانه خود قرار داده اند تقدیم نمایم.

از اساتید گرامیم جناب آقای دکتر حمید حسن پور، و جناب آقای دکتر وحید ابوالقاسمی بسیار سپاسگزارم چرا که بدون راهنمایی‌های ایشان تا این پایان نامه بسیار مشکل می‌نمود. همچنین از تمامی اساتید گرامی که تا این مقطع درسی مرا، بنمود فرمودند مراتب سپاسگزاری را دارم.

تعهد نامه

- اینجانب سجاد جعفری دانشجوی دوره کارشناسی ارشد رشته مهندسی کامپیوتر دانشکده مهندسی کامپیوتر دانشگاه صنعتی شاهرود نویسنده پایان نامه تحت راهنمایی 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 ارایه 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 بافتهای 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 ۱۳۹۷/۶/۱۸

دانشجو سجاد جعفری

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .

چکیده:

در این پایان‌نامه، به شناسایی و ردیابی عابرین پیاده در فریم‌های ویدیویی به صورت فشرده و زمان واقعی پرداخته می‌شود. یکی از چالش‌های اصلی ردیابی زمان واقعی، نیاز به حافظه زیاد است زیرا داده‌ها می‌بایست دائماً، به روزرسانی شوند. یکی از راه‌حل‌های این مساله انتخاب فریم کنونی و تعداد محدودی فریم بعدی است. مشکل این روش، کمبود داده‌های کافی برای آموزش است. در چنین مواقعی مدل آموزشی ضعیفی با داده‌های کم، شکل می‌گیرند که در شرایطی مثل تغییرات روشنایی، تغییرناگهانی زاویه هدف، تاری تصویر و انسداد مقاوم نیستند و ردیابی اشیاء را دچار رانش می‌کنند. تا زمانی که مدل آموزشی وابسته به داده‌ها باشند و بر این اساس، مدل‌ها آموزش ببینند، احتمال رانش و ردیابی اشتباه وجود دارد. در این پایان‌نامه از مدل غیرتطبیقی و از ساختار فضای ویژگی‌ها، برای آموزش استفاده شده است. یادگیری تقویتی نوعی از آموزش غیرنظارت‌شده و تلفیقی از روش آزمایش و خطا و پاداش می‌باشد. یادگیری تقویتی، همواره از چندین مسیر برای شناسایی و ردیابی هدف استفاده می‌کند و در مقابل انسداد، هوشمند است و می‌تواند فریم‌های از دست رفته را بازیابی کند و مجدداً به ردیابی هدف بپردازد. اما با توجه به همه این مسائل، احتمال رانش وجود دارد. در این پایان‌نامه از دو مرحله فشرده‌سازی بر روی داده‌های خام و فشرده‌سازی داده‌های تنک، استفاده شده است. فشرده‌سازی اولیه باعث می‌شود پیچیدگی محاسباتی کاهش یابد و در نتیجه صرفه‌جویی در مصرف حافظه صورت گیرد. برای اینکه رانش به حداقل مقدار ممکن برسد از فشرده‌سازی بر روی داده‌های تنک، استفاده شده است که همواره مناسب‌ترین داده‌ها را برای ردیابی انتخاب می‌کند. نتایج بدست آمده نشان می‌دهد الگوریتم پیشنهادی از کارایی بالایی در مقابل نویز و انسدادهای شدید برخوردار است به طوری که دقت و صحت ردیابی بر روی پایگاه داده^۱ به ترتیب، ۷۱/۵ درصد و ۳۶/۷ درصد بدست آمده است.

واژه‌های کلیدی: ردیابی شیء، یادگیری تقویتی، فشرده‌سازی، داده‌های تنک، مدل تطبیقی و غیرتطبیقی.

^۱https://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians.

۲۳ روش های تفریق زمینه
۲۳ بررسی الگوریتم های ACF, HOG, MDP
۲۴ الگوریتم HOG
۲۵ الگوریتم ACF
۲۶ الگوریتم MDP
۲۷ نتیجه گیری

۳. روش پیشنهادی

۳۰ مقدمه
۳۰ ساختار کلی روش پیشنهادی
۳۷ شناسایی و ردیابی هدف
۳۷ شناساگر عمومی و خاص یک شیء
۳۸ ردیابی هدف
۴۰ Haar-like
۴۱ الگوریتم فشرده سازی
۴۳ الگوریتم توسعه یافته تقویتی
۴۵ الگوریتم کلی ردیابی
۴۶ الگوریتم کلی ردیابی

۴. نتایج و آزمایش ها

۴۸ مقدمه
۴۸ شرایط انجام آزمایشات
۴۹ بررسی صحت تشخیص و ردیابی در مرحله آموزش و آزمایش
۵۲ بررسی تغییرات شیء از لحاظ معیارهای FN و FP
۵۳ مقایسه با مقاله پایه
۵۶ مقایسه با الگوریتم های پیشین
۵۷ درصد معیار منفی کاذب و مثبت حقیقی
۵۷ بررسی دقت و صحت اشیاء مورد ردیابی
۵۷ تعریف معیار دقت و صحت در ردیابی اشیاء
۵۸ مقایسه با مقاله پایه

- ۴-۵-۳- مقایسه با الگوریتم‌های پیشین ۶۰
- ۴-۶- بررسی MOTA با ویدیوهای تکی ۶۰
- ۴-۷- بررسی اجرای زمان الگوریتم پیشنهادی ۶۲
- ۴-۸- بررسی الگوریتم در صورت عدم فشردگی‌سازی در یکی از مراحل Lost و Tracking ۶۲
- ۴-۹- تاثیر فشردگی‌سازی‌های مختلف بر عملکرد الگوریتم ۶۴
- ۴-۱۰- نتیجه‌گیری ۶۶

۵. نتیجه‌گیری و کارهای آینده

- ۵-۱- جمع‌بندی و نتیجه‌گیری ۶۸
- ۵-۲- کارهای آینده ۷۰
- منابع ۷۱

فهرست اشکال

- شکل ۱-۱- بازشناسایی عابرین پیاده براساس حرکت. ۶
- شکل ۱-۲- نظارت خودکار ماشین‌ها در سطح آزادراه‌ها ۷
- شکل ۱-۳- ارتباط عامل با محیط از طریق آزمایش و خطا و پاداش در یادگیری تقویتی ۹
- شکل ۱-۲- تقسیم‌بندی الگوریتم‌های ردیابی براساس دویبعدی بودن ۱۸
- شکل ۲-۲- نقاط آشکارسازی شده توسط آشکارسازی‌های نقاط ۱۹
- شکل ۳-۲- ردیابی براساس هسته ۲۱
- شکل ۴-۲- ردیابی براساس سیاه‌نما ۲۲
- شکل ۵-۲- ردیابی براساس تفریق زمینه ۲۳
- شکل ۶-۲- نمونه اجرای الگوریتم HOG بر روی یک عکس در هر سلول ۲۶
- شکل ۷-۲- مراحل استخراج ویژگی در الگوریتم ACF ۲۵
- شکل ۸-۲- ردیابی چندعامله براساس الگوریتم مارکوف ۲۶
- شکل ۹-۲- فضای حالت الگوریتم مارکوف ۲۷
- شکل ۱-۳- نمایش خطی و غیرخطی بودن الگوریتم برای رسیدن به هدف ۳۱
- شکل ۲-۳- نمایش خطی و غیرخطی بودن مسیرها به هدف، از لحاظ بهینگی ۳۲
- شکل ۳-۳- نحوه اعمال فشرده‌سازی بر روی داده‌های خام و فشرده‌سازی بر روی ویژگی‌ها ۳۲
- شکل ۴-۳- نحوه اعمال فشرده‌سازی بر روی ویژگی‌ها در فرایند مارکوف ۳۳
- شکل ۵-۳- چارت روش پیشنهادی ۳۵
- شکل ۶-۳- روش ردیابی به وسیله شناسایی ۳۷
- شکل ۷-۳- نمونه‌ای از ردیابی یک شیء به وسیله الگوریتم جریان‌نوری ۳۸
- شکل ۸-۳- نمایش تصویر انتگرالی ۳۹
- شکل ۹-۳- نحوه قرارگیری نقاط در تصویر انتگرالی ۴۰
- شکل ۱-۴- چندنمونه از پایگاه‌داده‌های مورد ردیابی در پایان‌نامه ۴۷
- شکل ۲-۴- تشخیص و ردیابی اشیاء در مرحله آزمایش ۴۹
- شکل ۳-۴- تشخیص و ردیابی اشیاء در مرحله آموزش، در این مرحله شیء وارد حالت LOST می‌شود ۵۰
- شکل ۴-۴- تشخیص و ردیابی اشیاء در مرحله آزمون ۵۱
- شکل ۵-۴- نمایش برخی از فریم‌های مورد ردیابی ۵۲

فهرست جداول

- جدول ۴-۱ نتایج حاصل از FP و FN بر روی پایگاه داده‌هایی با انسداد شدید، براساس مقاله پایه..... ۵۳
- جدول ۴-۲ نتایج حاصل از FP و FN بر روی پایگاه داده‌هایی با انسداد شدید، براساس روش پیشنهادی..... ۵۳
- جدول ۴-۳ مقایسه نتایج حاصل از معیار منفی کاذب و منفی حقیقی با سایر روش‌های پیشین..... ۵۴
- جدول ۴-۴ نرخ معیار منفی حقیقی و منفی کاذب در روش پیشنهادی..... ۵۵
- جدول ۴-۵ نتایج حاصل از MOTA و MOTP بر روی پایگاه داده‌های با انسداد شدید، براساس مقاله پایه..... ۵۷
- جدول ۴-۶ نتایج حاصل از MOTA و MOTP بر روی پایگاه داده‌های با انسداد شدید، براساس روش پیشنهادی..... ۵۷
- جدول ۴-۷ نتایج حاصل از دو معیار دقت (MOTP) و صحت (MOTA) الگوریتم پیشنهادی با سایر روش‌های پیشین..... ۵۸
- جدول ۴-۸ نتایج حاصل از MOTA الگوریتم پایه..... ۵۹
- جدول ۴-۹ نتایج حاصل از MOTA بر اساس الگوریتم پیشنهادی..... ۵۹
- جدول ۴-۱۰ نتایج زمان اجرای الگوریتم پیشنهادی در مرحله آموزش و آزمایش..... ۶۰
- جدول ۴-۱۱ نتایج حاصل از MOTA و MOTP در صورت عدم فشرده‌سازی در حالت Tracking..... ۶۱
- جدول ۴-۱۲ نتایج حاصل از MOTA و MOTP در صورت عدم فشرده‌سازی در حالت Lost..... ۶۱
- جدول ۴-۱۳ نتایج حاصل از MOTA و MOTP با تعداد ویژگی ۴۰..... ۶۳
- جدول ۴-۱۴ نتایج حاصل از MOTA و MOTP با تعداد ویژگی ۶۰..... ۶۳

فصل اول

معرفی پژوهش

۱-۱. مقدمه

ردیابی اشیای متحرک یکی از مباحث چالش برانگیز در حوزه‌ی بینایی ماشین در سال‌های اخیر می‌باشد. آنالیز و فهم داده‌های موجود در تصویر و یافتن خودکار یک شیء مشابه در فریم‌های متوالی، از دلایل نیاز به ردیابی اشیاء متحرک است. یک سیستم ردیابی در کاربردهایی از جمله پاسخگویی سریع به رویدادها و مشاهده اشیاء، مورد نیاز می‌باشد. به همین دلیل از دهه پیش تاکنون توجه زیادی به موضوع ردیابی اشیاء متحرک شده است. به منظور آنالیز داده در صحنه‌ها و محیط‌های پیچیده می‌بایست عملیات ردیابی اشیاء را توسعه و گسترش داد که به این منظور تحقیقات گسترده‌ای در حوزه ردیابی اشیای متحرک صورت گرفته است. همچنین، ردیابی اشیاء متحرک به عنوان یک حوزه تحقیقاتی فعال در زمینه بینایی ماشین در سال‌های اخیر مطرح می‌باشد که کاربردهای گسترده‌ای در حوزه ربات‌های هوشمند، نظارت و مراقبت، تعامل بین انسان و کامپیوتر، ساختمان‌های هوشمند، ردیابی وسایل نقلیه و تصویربرداری پزشکی دارد.

سیستم‌های مراقبتی و نظارتی ویدیویی می‌بایست به محل نصب دوربین وابسته نباشند و همچنین این سیستم‌ها می‌بایست در مقابل تغییرات روشنایی محیط مقاوم باشند. به علاوه، این سیستم‌ها می‌بایست توانایی انطباق با تغییرات روشنایی ناشی از حرکت اشیاء را داشته باشند. تحقیقات اخیر تلاش زیادی برای توسعه سیستم ردیابی قابل تطبیق انجام داده‌اند به گونه‌ای که این سیستم‌ها به اندازه کافی برای کنترل تغییرات بیرونی همچون تغییرات روشنایی و انسداد اشیای چندگانه مفید واقع می‌شوند. بسیاری از کاربردهای ردیابی اشیای متحرک پیشنهاد شده به صورت خودکار حوادث اتفاق افتاده را تشخیص می‌دهند. به منظور تشخیص و درک بهتر این رخدادها نیازمند به انجام عملیات بینایی ماشین سطح پایین هستیم. از جمله این عملیات سطح پایین می‌توان به کنترل و مدیریت انسداد و تشخیص حرکات غیرمعمول اشاره کرد.

۱-۲. ویدیو

ویدیو، بیانگر دنباله‌ای پی‌درپی از تصاویر است که هر یک از این تصاویر، در لحظه‌ای از زمان نشان داده می‌شوند. به هر یک از این تصاویر ویدیویی، یک فریم می‌گویند [۱].

۱-۲-۱. نرخ فریم^۱

نرخ فریم یا تعداد فریم که آن را با^۲fps نمایش می‌دهند، نشان‌دهنده تعداد فریم‌های یک تصویر در یک ثانیه است. تعداد فریم‌های تصویر در ثانیه، زمانی اهمیت پیدا می‌کند که حرکتی در تصویر اتفاق بیافتد [۲]. امروزه به منظور کارایی بیشتر کیفیت تصویر، نرخ فریم در ویدیوها، ۶۰ فریم در ثانیه یا حتی بالاتر در نظر گرفته می‌شود اما در تلویزیون‌ها و سینما و اینترنت، نرخ فریم ۲۵ تا ۳۰ فریم در ثانیه، می‌باشد. در واقع نرخ فریم، برای کنترل سرعت جریان داده‌ها، بیان می‌شود [۳].

۱-۳. ردیابی^۳

هدف از ردیابی در تصاویر ویدئویی، یافتن موقعیت یک جسم یا یک شیء با توجه به مدل آموزشی است. ردیابی شیء کاربردهای فراوانی از جمله در کنترل ترافیک، تحلیل‌های ورزشی، دوربین‌های امنیتی در فرودگاه‌ها، فروشگاه‌ها و غیره دارد [۵، ۱].

در ردیابی اشیاء قبل از آن که فرآیند ردیابی آغاز شود، ابتدا نیاز به تشخیص شیء متحرک یا جداسازی آن از پس‌زمینه^۴ است [۱]. ردیابی هدف، یک مسئله دشوار است چون ظاهر مدل به دلیل عواملی چون روشنایی، انسداد و حرکت در میان دیگر شیء‌ها تغییر می‌کند. بنابراین ظاهر موثر مدل (مدل نمایشی)

^۱Frame rate

^۲Frame per second

^۳Tracking

یکی از مسائل مهمی است که در موفقیت یک الگوریتم ردیابی بسیار تاثیرگذار است و در سال‌های اخیر مورد توجه زیادی قرار گرفته است [۵،۸،۹،۱۰].

۱-۳-۱. رانش در ردیابی

عواملی مثل نویز، انسداد، روشنایی و تاری تصویر، ممکن است یک مدل نادرستی از هدف ارایه دهند که باعث یک تخمین نادرست از هدف شوند، به این عامل، رانش در ردیابی می‌گویند.

۱-۳-۲. ردیابی زمان واقعی و ردیابی آفلاین^۳

ردیابی شیء، می‌تواند به صورت زمان واقعی و یا آفلاین باشد. در ردیابی‌های آفلاین، نمونه‌ها از قبل توسط الگوریتم‌های با نظارت آموزش داده می‌شوند. در ردیابی‌های زمان واقعی، نمونه‌ها در طول زمان ردیابی، آموزش می‌یابند و غالباً به صورت بدون نظارت هستند. مزیت این روش، ردیابی سریع‌تر جسم هدف است [۴].

در ردیابی زمان واقعی، چون داده‌ها دائماً به روزرسانی می‌شوند در نتیجه فضای حافظه بسیار بالایی نیاز است و چنین فضایی در ردیابی‌های چند شیء غیرممکن است. همچنین باعث کندی در سرعت ردیابی می‌شود.

برای اینکه مشکل حافظه و سرعت در ردیابی زمان واقعی تا حدود زیادی حل شود، تنها چند فریم اخیر را برای ردیابی انتخاب می‌کنند و باقی فریم‌های بعدی را در نظر نمی‌گیرند. انتخاب اینکه تا چند فریم قبلی، در نظر گرفته شود، بسیار مهم است چون ممکن است باعث بروز رانش شود [۵].

¹ drift

² Online

³ Offline

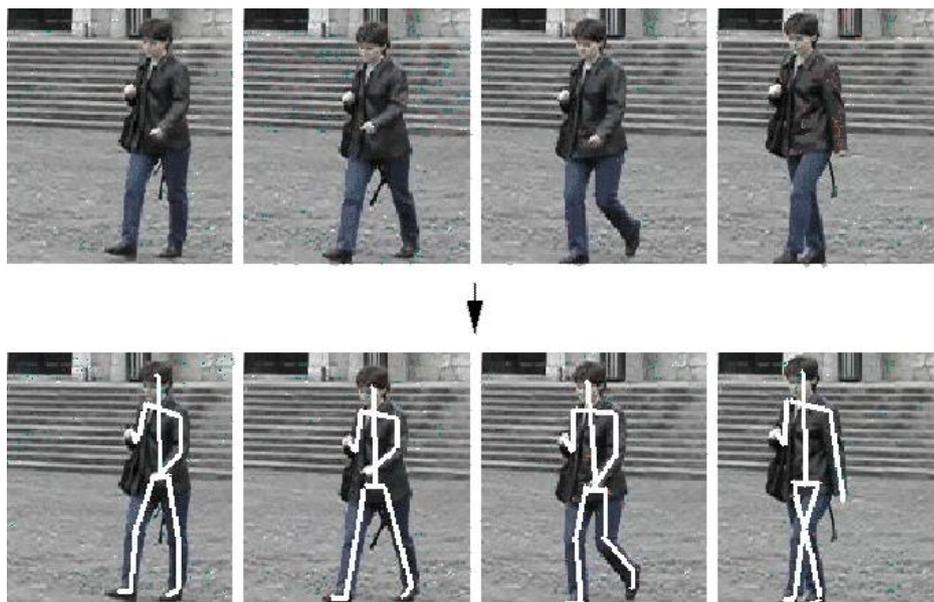
۱-۳-۳. برخی از کاربردهای ردیابی شیء

ردیابی شیء یک مرحله مقدماتی برای بسیاری از کاربردهای سطح بالای بینایی ماشین است و با گسترش کاربردهای بینایی ماشین در زندگی روزانه انسان هر روز بر اهمیت آن افزوده می‌شود. در ادامه چند کاربرد مهم ردیابی شیء معرفی شده است.

بازشناسایی براساس حرکت^۱

ردیابی شیء را می‌توان برای کاربردهای مختلفی مورد استفاده قرار داد که حرکات اشیاء کاملاً مختص به همان نوع شیء است. بدین ترتیب ما قادر خواهیم بود تا هویت و عملکرد اشیاء را از نوع حرکتشان تشخیص دهیم. به عنوان نمونه در مقاله [۶]، به شناسایی فرد براساس طرز راه رفتن، اشاره می‌کند در کاربردهای پزشکی هم از بازشناسایی حرکت، برای تحلیل حرکت توده‌های سرطانی یا تحلیل و بررسی نحوه‌ی راه رفتن بیماران مغز و اعصاب برای تشخیص روند بهبود یا پیشرفت بیماری به کار گرفته شده است. همچنین از بازشناسایی حرکت، می‌توان برای تعیین موقعیت افراد سالمند در منازل استفاده کرد تا اگر یک فرد سالمند به هر دلیلی مثلاً روی زمین بیافتد وضعیت او را تشخیص داده و به مراکز پزشکی اطلاع دهد [۷]. همچنین در کاربردهای ورزشی برای تحلیل حرکت بازیکنان در زمین مسابقه از ردیابی شیء می‌توان سود جست [۸]. بازشناسایی عابرین پیاده براساس مدل حرکتی در شکل ۱-۱ نشان داده شده است.

^۱Motion-based recognition



شکل ۱-۱- بازشناسایی عابرین پیاده براساس حرکت [۶].

در شکل ۱-۱، یک مدل حرکتی از هدف براساس آناتومی شیء، در نظر گرفته شده است. در هر فریم یک آناتومی از فرد ارایه شده و مجموعه‌ای از آناتومی‌های حرکتی، یک مدل برای تشخیص و ردیابی شیء ارایه خواهد داد.

نمایه‌گذاری فایل‌های ویدیویی

ردیابی شیء برای تفسیر و بازیابی خودکار ویدیو از میان مجموعه داده‌های ویدیویی بکار می‌رود. به طور نمونه می‌توان از شاخص‌گذاری^۱ و بازیابی ویدیو براساس محتوا در موتورهای جستجو نام برد [۹].

نظارت خودکار^۲

از ردیابی شیء می‌توان برای تشخیص صحنه‌های مشکوک یا رخداد‌های نامطلوب استفاده کرد. در فرودگاه‌ها انواع متفاوت کامیون‌ها باید در الگوهای متفاوتی حرکت کنند. همچنین ترکیبی از مکان‌ها و الگوها هستند که هرگز نباید اتفاق بیافتند. برای مثال کامیون نباید در مسیر فعال فرود توقف نماید،

¹ Indexing

² Automatic monitoring

سیستم‌های رایانه‌ای که بتواند این فعالیت‌ها را رصد نماید و اگر مسئله‌ای یافت شد هشدار دهد مفید خواهد بود [۱۰]. شکل ۱-۳، نمونه‌ای از نظارت خودکار ماشین‌ها در سطح آزادراه‌ها را نشان می‌دهد.



شکل ۱-۲- نظارت خودکار ماشین‌ها در سطح آزادراه‌ها [۱۰].

۴-۱. عامل‌ها^۱

در یک مسئله یادگیری، با عاملی روبرو هستیم که از طریق سعی و خطا با محیط تعامل کرده و یاد می‌گیرد تا عملی بهینه را برای رسیدن به هدف انتخاب نماید.

۱-۴-۱. یادگیری عامل‌ها^۲

یادگیری عامل‌ها می‌تواند به سه صورت، یادگیری نظارت‌شده یادگیری غیرنظارت‌شده و یادگیری تقویتی باشد. که در ادامه هر سه روش به طور خلاصه شرح داده می‌شود.

۱-۴-۲. یادگیری نظارت‌شده^۳

در این روش یک معلم یا ناظر وجود دارد که بهترین عمل را در هر وضعیت بلد است. در جمله‌ی، "اگر هوا ابری باشد، باید چتری به همراه داشته باشی." یک معلم یا ناظر وجود دارد که می‌گوید اگر هوا ابری

¹ agents

² Learning agents

³ Supervised learning

باشد بهترین عمل، به همراه داشتن چتر است. این عامل توصیه‌هایی برای تصحیح شیوه‌ی عملکرد ناظر ارایه می‌دهد [۱۳].

۱-۴-۳. یادگیری غیرنظارت‌شده^۱

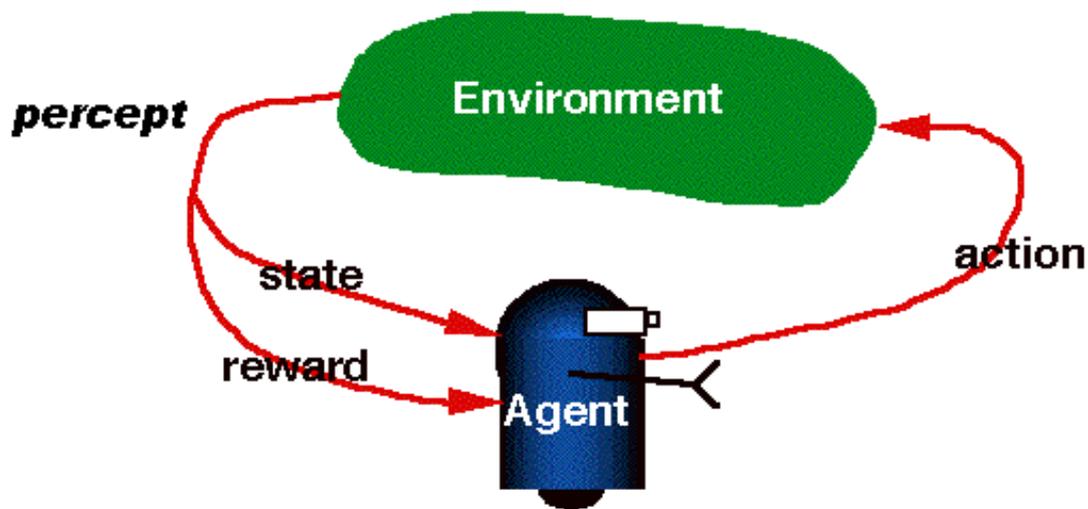
در این روش یک معلم یا ناظر وجود ندارد که بهترین عمل را در هر وضعیت بشناسد. در جمله‌ی "اگر هوا ابری باشد، و من چتری با خود نداشته باشم، به احتمال زیاد خیس خواهم شد." مشاهده‌ی نتیجه‌ی ترک کردن خانه، بدون چتر و در یک روز ابری، شکلی از فیدبک است. با این وجود، نتیجه‌های شبیه این، به طور مستقیم، پیشنهادی در مورد انتخاب اعمال بهتر ندارند. اینکه عامل ترجیح بدهد، که هنگام خروج از خانه چتری به همراه داشته باشد، کاملاً به این بستگی دارد که خیس شدن یا نشدن برای وی اهمیت داشته باشد [۱۳].

۱-۴-۴. یادگیری تقویتی^۲

در این نوع از یادگیری فیدبکی به صورت عبارات کمکی مثبت (پاداش) و یا منفی (جریمه) به عامل یادگیرنده داده می‌شود. در جمله‌ی، "سابقاً هوا ابری بود و تو چتری به همراه نداشتی؛ در نتیجه تو خیس شدی، و این خیلی بد است." عامل یاد می‌گیرد اگر هوا ابری باشد و چتری به همراه نداشته باشد خیس می‌شود در نتیجه جریمه می‌شود. در شکل ۱-۶، عامل با محیط در تعامل است و با توجه به عملی که انجام می‌دهد، پاداش می‌گیرد و این پاداش منفی یا مثبت به ازای هر عمل، باعث یادگیری عامل در تعامل با محیط می‌گردد.

¹ Non- Supervised learning

² Reinforcement learning



شکل ۱-۳- ارتباط عامل با محیط از طریق آزمایش و خطا و پاداش در یادگیری تقویتی [۱۳].

۱-۵. فشردن سازی^۱

فشردن سازی داده‌ها، به معنای کاهش میزان داده برای انتقال مقدار مشخصی از اطلاعات است. فشردن-سازی براساس نوع داده، به دو دسته زیر تقسیم می‌شوند:

الف)- کاهش افزونگی^۲ بر روی داده‌های خام

ب)- داده‌های تنک^۳ بر روی ویژگی‌ها [۵، ۱۵، ۱۶].

در قسمت الف) از شیوه‌های متفاوتی برای توصیف مقدار مشخصی از اطلاعات می‌توان استفاده کرد و ممکن است هر یک میزان مختلفی از داده‌ها را به کار گیرند. برخی از این توصیف‌ها ممکن است، حاوی اطلاعات نامرتب یا تکراری باشند که در واقع بیانگر افزونگی هستند. هدف از این نوع فشردن سازی، کم کردن حجم داده با کاهش یا حذف افزونگی‌های موجود در آن است.

^۱ Compression

^۲Reduce redundancy

^۳Sparse

اطلاعات نامرتبط

اطلاعات نامرتبط موجود در یک تصویر، اغلب به عنوان افزونگی مورد توجه قرار می‌گیرند، و به دو جنبه افزونگی برای کاربردی خاص و افزونگی بصری اشاره دارند. افزونگی بصری به این اشاره دارد که همه نواحی تصویر از نظر سیستم بینایی انسان به یک میزان اهمیت ندارند و برخی از اطلاعات تصویر در سیستم بینایی انسان نادیده گرفته می‌شوند. از این رو این اطلاعات افزونه را می‌توان از تصویر حذف کرد، بدون آن که توسط چشم انسان قابل درک باشد [۱].

در قسمت ب) بر روی ویژگی‌ها، فشرده‌سازی انجام می‌گیرد و ویژگی‌هایی انتخاب می‌شوند که به نسبت بقیه از اهمیت بالاتری برخوردار هستند.

قاعدتا انتظار می‌رود که فشرده‌سازی بر روی داده‌های خام بیشتر در جهت صرفه‌جویی در حافظه و بهبود سرعت، کمک‌کننده باشند. در فشرده‌سازی داده‌های تنک، انتظار می‌رود با انتخاب بهترین ویژگی‌ها، دقت ردیابی افزایش یابد و رانش به حداقل برسد [۵].

۱-۶. تعداد فریم‌ها در هر مرحله آموزش

در ردیابی‌های زمان واقعی، به علت اینکه اطلاعات، دائما بروزرسانی می‌شوند نیاز به فضای حافظه بسیار زیادی می‌باشد. عملا نمی‌توان از اطلاعات همه‌ی فریم‌ها برای شناسایی و ردیابی اشیاء استفاده کرد. برای اینکه پیچیدگی محاسباتی کم شود و فضای حافظه کاهش یابد، تنها چند فریم اخیر برای ردیابی انتخاب و باقی فریم‌ها کنار گذاشته می‌شوند. در این پایان‌نامه از ۱۰ فریم اخیر برای شناسایی و ردیابی استفاده شده است و فریم‌های قبل آن، کنار گذاشته شده‌اند.

۱-۷. تعریف مسئله

در الگوریتم‌های ردیابی زمان واقعی، پردازش اطلاعات زیاد است، در نتیجه پایگاه داده بزرگ نیاز دارد در نتیجه حجم داده‌ها افزایش می‌یابد و نتیجتاً پیچیدگی محاسباتی بالا می‌رود. انتخاب الگوریتم‌های مبتنی بر فشرده‌سازی حجم داده‌ها را کاهش می‌دهد [۴].

الگوریتم‌های ردیابی زمان واقعی متعددی وجود دارند که اغلب مدل‌ها در آن با توجه به نمونه مشاهدات فریم‌های اخیر به روزرسانی می‌شوند. اما در آن‌ها مشکل رانش یک مسئله جدی و مهم است [۵].

روش‌های پیشین تا حدی توانسته مشکل رانش را در ردیابی شیء هدف برطرف کنند اما همچنان تحقیقات برای بهبود این مسئله در جریان است. همچنین مدل‌های فعلی آرایه شده برای انتخاب نمونه‌ها در راستای تشخیص هدف، برای آموزش، دقیق نبوده و باعث ایجاد یک مدل نمایشی ضعیفی می‌شوند. علاوه بر این، با توجه به ضرورت ردیابی زمان واقعی در کاربردهای امروزی، آرایه دادن یک مدل دقیق با حداقل خطای ممکن برای شناسایی دقیق هدف (عابرین پیاده) لازم و ضروری است. از طرفی با توجه به بالا بودن حجم داده، آرایه مدلی با حداقل پیچیدگی محاسباتی و پردازش سریع داده‌ها بر ضرورت تحقیق در این زمینه صحنه می‌گذارد.

۱-۸. هدف و نوآوری پایان‌نامه

برای اینکه مشکل حافظه و سرعت در ردیابی زمان واقعی تا حدود زیادی حل شود، تنها چند فریم اخیر برای ردیابی انتخاب می‌شود و باقی فریم‌های قبلی در نظر گرفته نمی‌شود. چگونگی انتخاب تعداد فریم‌های بعدی بسیار مهم است، چون ممکن است ردیابی باعث بروز رانش شود [۴، ۵، ۱۸].

در ردیابی زمان واقعی با مشکل حافظه و سرعت روبرو است برای حل این مسائل، تنها چند فریم اخیر را در نظر می‌گیریم. اما با همه‌ی این تفصیلات، مشکل حافظه و سرعت بطور کامل بهبود نیافته است. همچنین به علت انتخاب محدودی از فریم‌های اخیر، ممکن است ردیابی دچار رانش شود. کاهش افزونگی

یعنی کوچک شدن اندازه فریم‌ها به طوری که اطلاعات از بین نرود. می‌توان با استفاده از فشرده‌سازی افزونگی، بخشی از اطلاعات پس‌زمینه را حذف کرد در نتیجه مشکل حافظه و سرعت بسیار بهبود می‌یابد. اما همچنان مشکل رانش باقی می‌ماند. در استخراج ویژگی‌ها، می‌توان ویژگی‌هایی را انتخاب کرد که بیشتر در مقابل نویز مقاوم باشد، این کار توسط فشرده‌سازی داده‌های تنک انجام می‌شود که هم باعث می‌شود مشکل رانش به حداقل برسد هم باعث بهبود سرعت و حافظه شود.

با ترکیب یادگیری تقویتی در مرحله انتخاب ویژگی‌های داده‌های تنک، می‌توان پیچیدگی محاسباتی و رانش را به حداقل برسد، یادگیری تقویتی با اعمال سیاست، بهترین دیکشنری را در کمترین زمان ممکن با هرس کردن بخش‌های زیادی از ویژگی‌های نامرتبب انجام می‌دهد. در واقع هدف ما این است که تحت قالب داده‌های تنک، سیاست تقویتی بر روی داده‌ها اعمال کنیم [۵، ۱۷، ۱۸].

هدف ما ارایه یا بهبود یک الگوریتم ردیابی زمان واقعی برای ردیابی عابرین پیاده می‌باشد. به طوری که مشکل رانش را به حداقل برسانیم.

۹-۱. ساختار پایان‌نامه

این پایان‌نامه در ۵ فصل کلی تدوین شده است. که در ابتدای هر فصل مروری بر مطالب آن صورت خواهد گرفت. در فصل اول، مقدمه و ادبیات موضوع بیان شد. در فصل دوم، پیشینه‌ی تحقیق و دسته‌بندی از متدوال‌ترین روش‌های حل مسئله‌ی ردیابی شیء مرور و ارایه خواهد شد. و در انتها پیچیدگی‌ها و موانع مسئله‌ی ردیابی شیء بیان خواهد شد. در فصل سوم، الگوریتم پیشنهادی ارایه خواهد شد. در فصل چهارم، نتایج الگوریتم پیشنهادی بررسی خواهد شد و در فصل پایانی نتیجه‌گیری و پیشنهادهایی برای کارهای آینده ارایه خواهد شد.

فصل دوم

مبانی نظری و پیشینه پژوهش

۲-۱. مقدمه

در این فصل پیشنهادی تحقیق و روش‌هایی که در حوزه ردیابی اشیاء هستند بررسی می‌شود. الگوریتم‌های ردیابی را می‌توان از جهات مختلف، براساس مدل، دوبعدی یا سه‌بعدی بودن شیء‌ها تقسیم‌بندی کرد. و در آخر به بررسی سه الگوریتم ACF, HOG, MDP که برای شناسایی اشیاء قبل از ردیابی استفاده می‌شوند می‌پردازیم.

۲-۲. تقسیم‌بندی الگوریتم‌های ردیابی براساس مدل

الگوریتم‌های ردیابی براساس مدل به ۳ دسته زیر تقسیم می‌شوند [۱۹، ۲۰، ۲۱]:

۱- مدل حرکتی

۲- مدل هندسی

۳- مدل ظاهری

سیستم‌هایی که در سال‌های اخیر برای ردیابی استفاده شدند عموماً از سه نوع ویژگی مدل حرکت، مدل هندسی و مدل ظاهر افراد استفاده کردند هر یک از این ویژگی‌ها می‌توانند عملکرد متفاوتی در ردیابی و انطباق افراد در زمان ورود به زوایه دید دوربین داشته باشند [۱۴].

۲-۲-۱. مدل حرکتی

در ردیابی بر پایه حرکت پنجره از شیء، فریم فعلی که پارامترهای حرکتش مانند مکان، سرعت و شتاب شبیه پارامترهای حرکت فریم قبلی هستند تقسیم‌بندی می‌شود. در این نوع ردیابی‌ها از جرم هندسی افراد که به آن، مرکز ثقل نیز گفته می‌شود، استفاده می‌شود. مرکز ثقل ویژگی ثابتی می‌باشد که اگر تاخیر بسیار کمی در شکل افراد به وجود بیاید تاثیر چندانی در آن ایجاد نمی‌شود [۱۴].

از مهمترین روش‌های مبتنی بر حرکت می‌توان به فیلترکالمن^۱، فیلترذرات^۲ و جریان نوری^۳ اشاره کرد. الگوریتم فیلترذره شبیه فیلترکالمن عمل می‌کند اما یک سری مزیت‌هایی نسبت به فیلترکالمن دارد که عبارتند از: ۱- الگوریتم فیلترذره در هر مرحله به روزرسانی می‌شود و از میان ذرات موجود بهترین ذرات را برای حالت بعدی انتخاب می‌کند که انتخاب بهترین ذرات با استفاده از الگوریتم likelihood انجام می‌شود [۱۸]. یکی از معمول‌ترین روش‌ها برای تخمین حرکت در تصاویر دوبعدی جریان نوری است. جریان نوری عبارت است از حرکت الگوی روشنایی در دنباله‌ای از تصاویر، که می‌تواند تقریبی از حرکت میزان هر پیکسل یا گروهی از پیکسل‌های موجود در یک تصویر باشد [۲۲].

در [۱۸] براساس داده‌های تنک و سنجش فشرده و درجه تنکی و براساس مدل حرکتی فیلترذرات به ردیابی افراد می‌پردازد. در ردیابی برپایه حرکت مانند فیلترکالمن [۲۳] از توانایی پیش‌بینی مکان افراد در فریم بعدی با استفاده از تاریخچه حرکت آن استفاده می‌شود. در [۲۴] یک روش جدید فیلترذره‌ای بادوام معرفی شده، که این روش با نمونه‌گیری آماری برای کاهش تعداد نمونه‌های مورد نیاز ترکیب می‌شود. مشکل اصلی در ردیابی انسان با فیلترذره‌ای، رشد نمایی ذرات به منظور تخمین صحیح مسیر حرکت افراد می‌باشد [۲۵، ۲۶]. در [۱۷] از روش جریان نوری و براساس تاریخچه هدف، ردیابی انجام شده است. در این مقاله از طریق Forward-Backward همپوشان بودن و نبودن نقاط را بررسی می‌کند. در [۶۲] از مدل تصمیم‌گیری فازی با استفاده از فیلترذره پرداخته شده است که به ردیابی عابرین پیاده در انسدادهای جزئی می‌پردازد. در [۶۳] از الگوریتم ردیابی پیاده‌سازی پیشرفته با توالی‌های تصویر ارائه شده است، توالی‌های تصویر توسط دوربین‌های نظارتی به دست می‌آید.

¹ Calman Filter

² Particle filter

³ Optical current

۲-۲-۲. مدل هندسی

در مدل هندسی از ویژگی‌های شکل برای شناسایی و ردیابی استفاده شده است. تشخیص لبه، تطبیق مرز شیء، گشتاور، مساحت و اندازه از جمله ویژگی‌های شیء محسوب می‌شوند که برای ردیابی مورد استفاده قرار می‌گیرند. در [۲۷] براساس الگوریتم لبه‌گیری کنی، عمل می‌کند پس از تفریق پیش‌زمینه از پس‌زمینه و حذف نویز از تصویر حاصل، با استفاده از لبه‌گیری کنی، لبه‌های نواحی پیش‌زمینه بدست می‌آید. در نهایت با استفاده از احتمال گوسی، اشخاص با توجه به جایگاهشان در فرم قبلی ردیابی می‌شوند. در [۶۴] برای حل مشکل ترافیک عابرپیاده و ایمنی عابرپیاده در محیط ترافیکی شهری، روش تشخیص و ردیابی عابرپیاده براساس دید مونوکولار مورد مطالعه قرار گرفته است.

۳-۲-۲. مدل ظاهری

اطلاعات ظاهر اشیاء به مکرر در شناسایی و ردیابی استفاده شده است. از جمله مهم‌ترین روش‌های موجود برای شناسایی افراد با استفاده از ظاهر آن‌ها می‌توان به تابع چگالی احتمال، مدل ظاهری یک-بعدی و دوبعدی، چندین نقطه ثابت در افراد و مدل ظاهری از چند زاویه دید، اشاره کرد.

در [۱۸] از الگوریتم RTCST استفاده شده است. برای کاهش ابعاد و بهینه‌سازی، از الگوریتم OMP^۱ استفاده شده است. با استفاده از الگوریتم OMP سرعت ردیابی ۵ هزار برابر سریع‌تر به نسبت الگوریتم‌های چون L1Tracker گردیده است.

در [۲۸] روش جدیدی با استفاده از الگوریتم خوشه‌بندی K-means اتوماتیک که اطلاعات قبلی در مورد تعداد اشیاء متحرک ندارد برای ردیابی اشیاء متحرک ارائه می‌کند. ایده اصلی این روش این است که

^۱ Real-time Compressive Sensing Tracking

^۲ Orthogonal Matching Pursuit

ابتدا ویژگی‌های شیء و زمینه نزدیک به شیء را خوشه‌بندی کرده و سپس با حذف خوشه‌های مشترک، شیء را از زمینه اطرافش متمایز کرده و ردیابی می‌کند.

تقسیم‌بندی الگوریتم‌های ردیابی براساس مدل ظاهری

الگوریتم‌های ردیابی براساس مدل نمایشی به طور کلی به دو دسته تقسیم می‌شوند:

۱-مدل مولد [۵،۲۹،۳۰،۳۱]

۲-مدل متمایزکننده [۵،۲۹،۳۰،۳۱]

مدل مولد

در الگوریتم ردیابی مولد، یک مدل برای نشان دادن تصویر یا شیء هدف آموزش می‌بیند و سپس مدل آموزش دیده برای جستجوی ناحیه هدف با حداقل بازسازی خطا عمل می‌کند.

اخیرا روش‌های مبتنی بر داده‌های تنک یک مدل L1-tracker [۵،۱۵] ارائه نموده است که در آن جسم با توجه به ترکیب خطی از داده‌های تنک هدف و قالب‌های بی اهمیت، مدل می‌شوند. اما پیچیدگی محاسباتی آن بالاست، در نتیجه استفاده از آن‌ها در دنیای واقعی محدودیت وجود دارد. برای گسترش آن و حل این مشکل از الگوریتم ردیابی متعامد [۴] استفاده شده است.

مدل متمایزکننده

الگوریتم‌های متمایزکننده روش‌هایی هستند که مسئله ردیابی در آن به عنوان یک طبقه‌بندی دودویی عمل می‌کند و وظیفه آن پیدا کردن مرز تصمیم‌گیری برای جدا کردن شیء هدف از پس‌زمینه است.

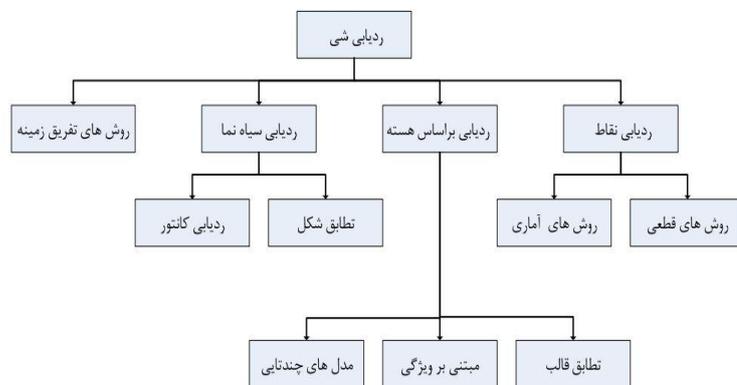
¹ Generative

² Discriminative

اما الگوریتم های متمایز کننده هم به علت نويز و نمونه های منحرف اغلب منجر به رانش می شوند. در سال های اخیر از یک روش نیمه نظارت^۱ [۵،۳۲] تقویت زمان واقعی برای کاهش مشکل رانش استفاده شده است که در آن تنها نمونه ها در فریم اول با برچسب و همه نمونه های دیگر بدون برچسب هستند.

۲-۳. تقسیم بندی الگوریتم های ردیابی به طور کلی براساس دوعدی بودن

الگوریتم های ردیابی براساس رویکرد دوعدی بودن به صورت زیر تقسیم بندی می شوند:



شکل ۲-۱ تقسیم بندی الگوریتم های ردیابی براساس دوعدی بودن

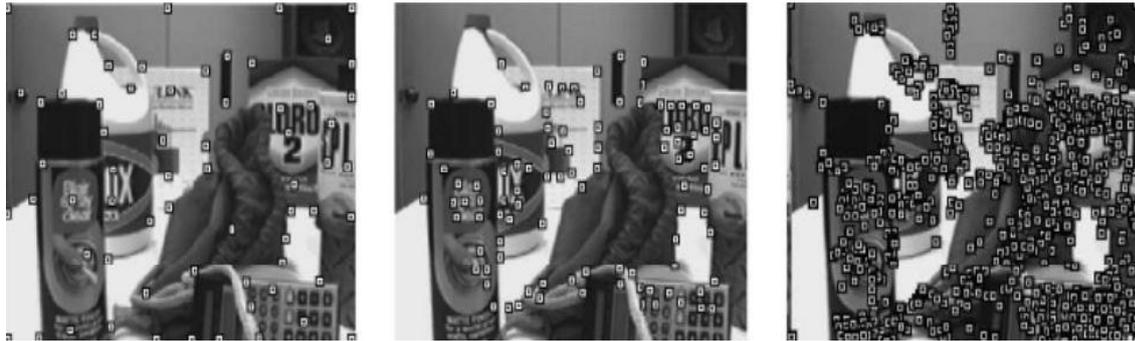
۲-۳-۱. ردیابی نقاط

ردیابی می تواند بصورت تطبیق اشیاء بازنمایی شده با نقاط در طول فریم ها فرموله شوند. این نقاط ابتدا توسط آشکارسازهای نقاط مشخص می گردند. آشکارسازهای نقاط برای یافتن نقاطی در تصویر استفاده می شوند که دارای بافت های پرمعنایی در مکان هایشان باشند.

آشکارسازهای نقاط بکار رفته، شامل KLT, Harris, Moravec, SIFT است. یک ارزیابی مقایسه ای از آشکارسازهای نقاط در [۳۳،۳۴] ارائه شده است. مثالی از آشکارساز نقاط SIFT, KLT, Harris در شکل ۲-۲، نشان داده شده است. به طور کلی آشکارسازی نقاط SIFT تعداد نقاط بیشتری در مقایسه با

^۱Semi-supervised

دیگر آشکارسازی‌های نقاط تولید می‌کند. دلیل این مساله این است که در این رویکرد، نقاط در مقیاس‌ها و رزولوشن‌های مختلف باهم جمع شده‌اند. بطور تجربی در [۳۳،۳۴] نشان داده شده است که STFT از اکثر آشکارسازی‌های نقاط نتایج بهتر و مفیدتری دارد و در برابر تغییر شکل‌های تصویر انعطاف‌پذیری بیشتری دارد.



شکل ۲-۲ نقاط آشکارسازی‌شده توسط آشکارسازی‌های نقاط، به ترتیب از سمت چپ نقاط آشکارساز Harris، نقاط آشکارساز KLT، نقاط آشکارساز SIFT [۳۳،۳۴].

الگوریتم‌های ردیابی نقاط را الگوریتم‌های ردیابی مبتنی بر تطبیق مشخصه نیز می‌نامند. به دلیل ناتوانی در جداسازی پیش‌زمینه از پس‌زمینه، الگوریتم‌های ردیابی مبتنی بر تطبیق مشخصه در ردیابی اشیاء صلب نتایج بهتری را ارائه می‌دهند؛ چون هنگامی که شیء غیرصلب باشد، ویژگی‌های استخراج شده از پس‌زمینه الگوریتم را با مشکل مواجه می‌کند [۳۵].

به طور کلی روش‌های تطبیق نقاط می‌تواند به دو دسته بزرگ زیر تقسیم شوند:

۱- روش‌های قطعی

۲- روش‌های آماری

روش‌های قطعی

روش‌های قطعی برای تطبیق نقاط با استفاده از مجموعه‌ای از شروط حرکتی، یک هزینه وابستگی برای هر شیء در فریم $t-1$ به یک شیء در فریم t تعریف می‌کنند [۳۶].

در [۳۷] نویسندگان مسئله تطابق نقاط را با رویکرد حریصانه و براساس شروط مجاورت و صلب بودن حل کردند. در [۳۸] یک رویکرد حریصانه که با تخمین یکنواخت مقید شده ارایه شده است.

روش‌های آماری

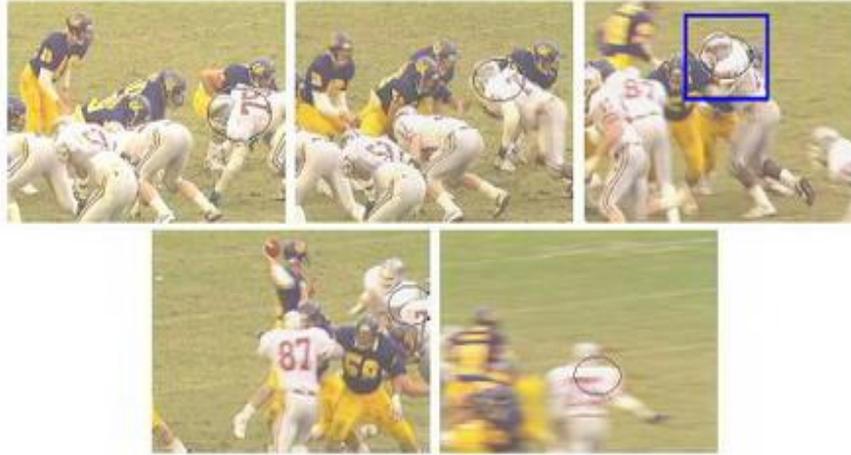
اندازه‌گیری‌های انجام شده از دنباله‌های ویدیویی همواره شامل نویز است. در روش‌های تخمین آماری مانند فیلترکالمن و فیلتر ذره‌ای از رویکرد فضای حالت برای مدل کردن ویژگی‌های خاصی از شیء مانند موقعیت، سرعت و شتاب استفاده می‌کنند [۳۹]. در [۴۰] یک الگوریتم استخراج جسم حرکتی مبتنی بر ردیابی را ارایه شده است که در آن مکان و ردیابی هدف با استفاده از یک روش تطبیق آماری به دست می‌آید. مناطق ساختاری اشیاء متحرک با استفاده از تکنیک تقسیم‌بندی مبتنی بر آستانه‌ها و به طور گسترده‌ای در وضوح بلوک نشان داده شده‌اند.

۲-۳-۲. ردیابی براساس هسته^۱

ردیابی بر پایه هسته در تقسیم‌بندی الگوریتم‌های ردیابی دوبعدی به عنوان یک شاخه اصلی در نظر گرفته می‌شود. شکل ۲-۳، ردیابی براساس هسته را نشان می‌دهد که در آن یک شیء به وسیله یک مستطیل مدل می‌گردد. در [۴۱] یک روش ردیابی اشیاء غیرصلب به کمک تکرار جابجایی میانگین مبتنی بر هسته یکنوا کاهشی و به کمک حذف پس‌زمینه ارایه می‌شود.

^۱KernelTracking

الگوریتم‌های ردیابی براساس هسته را می‌توان به سه دسته، ردیابی براساس تطابق قالب، براساس ویژگی، و براساس چند زاویه دید تقسیم کرد.



شکل ۲-۳ ردیابی براساس هسته [۴۱].

ردیابی براساس تطابق قالب

تطابق قالب یک رویکرد جستجوی سراسری در تصویر برای یافتن یک ناحیه‌ی مشابه با قالب شیء است. قالب‌ها توسط شکل‌های هندسی مثل بیضی یا مستطیل ساخته می‌شوند. در [۴۲] یک الگوریتم کارآمد براساس تطابق قالب ارائه شده است.

روش‌های مبتنی بر ویژگی

در روش‌های مبتنی بر ویژگی، از ویژگی‌های شیء همچون رنگ، بافت، عمق شیء و کنتراست استفاده می‌کنند. در [۴۳] نویسندگان، مدل شیء را توسط میانگین رنگ و پیکسل‌های درون ناحیه مستطیلی ساختند و برای کاهش محاسباتی در فریم بعد شیء را در ۸ همسایگی اطراف شیء جستجو کردند. در

^۱Texture
^۲Contrast

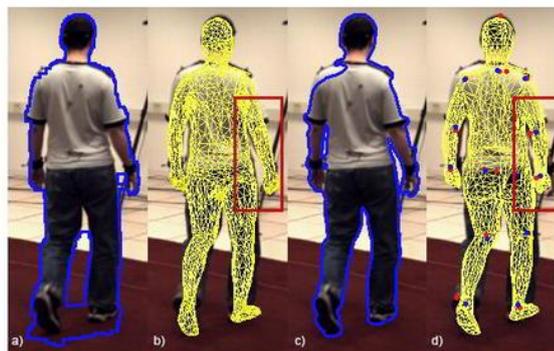
[۴۴] از هیستوگرام وزن داده شده در ناحیه بیضوی برای بازنمایی شیء استفاده کردند و بجای انجام جستجوی سراسری آن‌ها از پروسه‌ی جابجایی میانگین در [۴۵،۴۶] برای مکان‌یابی شیء استفاده کردند.

روش‌های مدل‌های چندتایی

در روش‌های مبتنی بر هسته که مدل شیء توسط قالب، هیستوگرام و غیره بیان می‌شود معمولاً مدل شیء به صورت بهنگام ساخته می‌شود بنابراین این مدل‌ها اطلاعات بدست آمده از آخرین مشاهدات شیء را بازنمایی می‌کنند. برای غلبه بر این مشکل نماهای مختلف از شیء از قبل می‌تواند به صورت آفلاین به یک شبکه آموزش‌پذیر یاد داده شوند و برای ردیابی مورد استفاده قرار گیرند [۴۷]. در [۴۸] از طبقه‌بندی‌های ماشین‌های بردار پشتیبان برای ردیابی استفاده کرده است.

۲-۳-۳. ردیابی سیاه‌نما^۱

اشیاء ممکن است شکل‌های پیچیده‌ای داشته باشند برای مثال دست‌ها، سر و شانه‌ها و بدن انسان‌ها نمی‌توانند با شکل‌های ساده هندسی توصیف گردند. در شکل ۲-۴، ردیابی براساس سیاه‌نما نشان داده شده است در این شکل، شیء توسط محدوده‌ی دقیق آن مدل می‌گردد.



شکل ۲-۴ ردیابی براساس سیاه‌نما [۴۹].

^۱Silhouette tracking

در [۴۹] از ترکیب تفکیک پس‌زمینه تطبیقی و مدل مبتنی بر سیاه‌نمای آموزش دیده برای تشخیص و فیلترکالمن برای ردیابی هدف استفاده می‌کند.

۲-۳-۴. روش‌های تفریق زمینه^۱

در روش مدل‌سازی پس‌زمینه، از یک مدل آماری که توصیف‌کننده احتمالی مقادیر پس‌زمینه است، برای جدا کردن ناحیه پیش‌زمینه از ناحیه پس‌زمینه استفاده می‌شود. شکل ۲-۵، ردیابی براساس تفریق پس‌زمینه است. در [۵۰] با استفاده از روش تفریق پس‌زمینه و عملیات مورفولوژی وسایل نقلیه عبوری را تشخیص می‌دهد.



شکل ۲-۵ ردیابی براساس تفریق زمینه [۵۰].

در [۵۱] به طراحی و پیاده‌سازی سیستم ردیابی انسان به روش تفاضل پس‌زمینه پرداخته شده است.

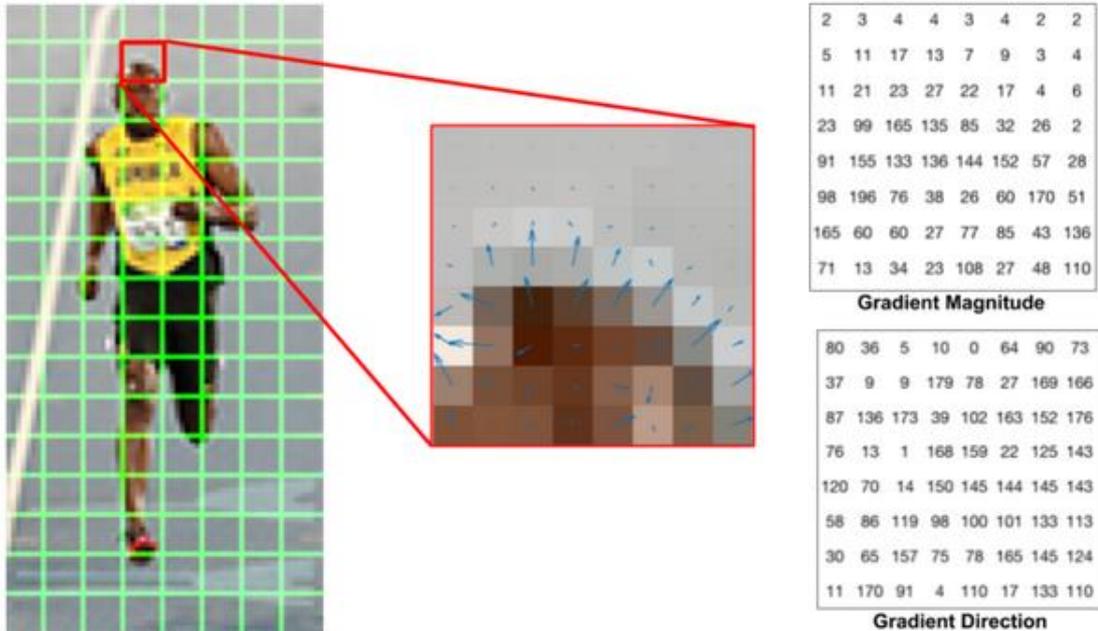
۲-۴. بررسی الگوریتم‌های ACF, HOG, MDP

برای تخمین مسیر یک شیء، ابتدا باید هدف را از طریق شناساگر شناسایی کرد و سپس به ردیابی شیء پرداخت. در زیر به تشریح سه الگوریتم معروف در این زمینه پرداخته می‌شود.

^۱Background subtraction

۲-۴-۱. الگوریتم^۱ HOG

روش هیستوگرام گرادیان‌های جهت‌دار در سال ۲۰۰۵ توسط Dalal و Triggs برای شناسایی انسان معرفی شد و تاکنون در کاربردهایی از جمله تشخیص و بازشناسی اشیاء بکار رفته است [۵۲].



شکل ۲-۶ نمونه اجرای الگوریتم HoG بر روی یک عکس در هر سلول [۵۵].

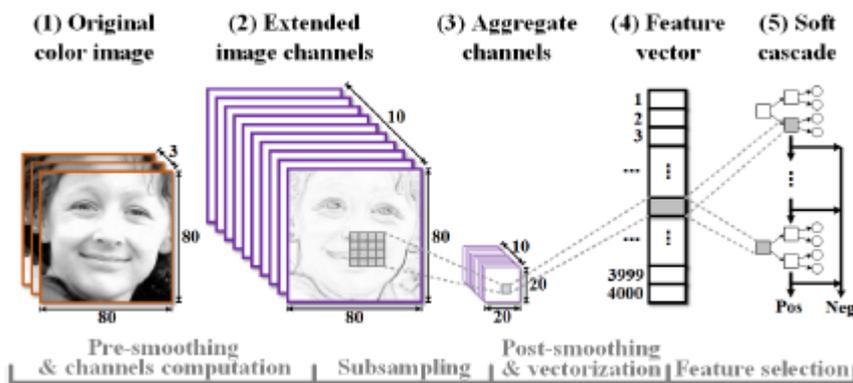
شکل ۲-۶، نمونه اجرای الگوریتم HoG بر روی یک عکس در هر سلول را نشان می‌دهد که هیستوگرام جهت‌های گرادیان برای پیکسل‌های داخل سلول محاسبه می‌شود. سپس برای مقاومت در برابر تغییرات روشنایی، هیستوگرام چند سلول مجاور که با هم تشکیل یک بلوک را می‌دهند، نرمالیزه می‌شود. این هیستوگرام نرمالیزه شده، هیستوگرام گرادیان‌های جهت‌دار نامیده می‌شود. برای تصاویر رنگی، گرادیان برای هر کانال رنگی بصورت جداگانه محاسبه می‌شود و بزرگترین مقدار برای هر پیکسل بعنوان بردار گرادیان آن پیکسل انتخاب می‌شود [۵۵].

^۱ Histogram of oriented gradients

۲-۴-۲. الگوریتم ACF

الگوریتم ACF از مفهومی به نام کانال برای استخراج ویژگی‌ها استفاده می‌کند. این کانال‌ها انواع مختلفی مثل کانال اندازه شیب، کانال هیستوگرام شیب‌گرا می‌توانند داشته باشند.

مراحل الگوریتم ACF برای استخراج ویژگی‌ها در شکل ۲-۷، نشان داده شده است [۵۶].



شکل ۲-۷ مراحل استخراج ویژگی در الگوریتم ACF، [۵۶].

همان‌طور که در شکل ۲-۷، نشان داده شده، اساس این الگوریتم بر پایه کانال است. یکی از مهم‌ترین نوع کانال‌ها، کانال رنگ است که مهم‌ترین نوع رنگ‌های به کار برده شده Gray-Scale و RGB است. برای اینکه الگوریتم ACF بهترین عملکرد را داشته باشد باید نوع کانال مورد استفاده، اندازه پنجره، روش‌های نمونه برداری و اندازه ویژگی به درستی تعیین شود. سه نوع کانالی که در ACF مورد استفاده قرار می‌گیرد شامل کانال رنگ (RGB, Gray-Scale, HSV)، کانال اندازه شیب و هیستوگرام شیب می‌باشند.

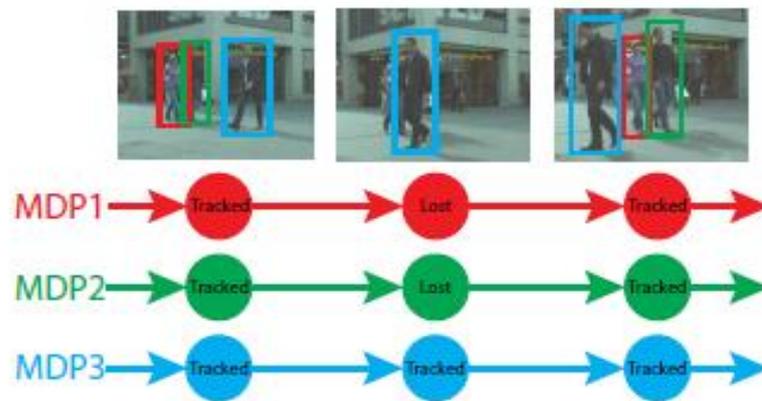
\gradient magnitude

\oriented gradient histograms

اندازه پنجره، مقیاسی برای انتخاب تغییر اندازه نمونه‌های هدف و غیرهدف و سپس آموزش پیکسل-هایبست که انتخاب شدند.

۲-۳-۴. الگوریتم^۲ MDP

یکی از روش‌های یادگیری تقویتی، روش مارکوف است که برای استخراج ویژگی داده‌ها استفاده می‌شود. شکل ۲-۸، ردیابی چندعامله براساس الگوریتم مارکوف را نشان می‌دهد. که در آن هر شیء به صورت یک فرایند مارکوف به صورت جداگانه مدل می‌گردد. در واقع، در فرایندهای مارکوف طول عمر^۳ یک شیء با یک MDP مدل می‌گردد.



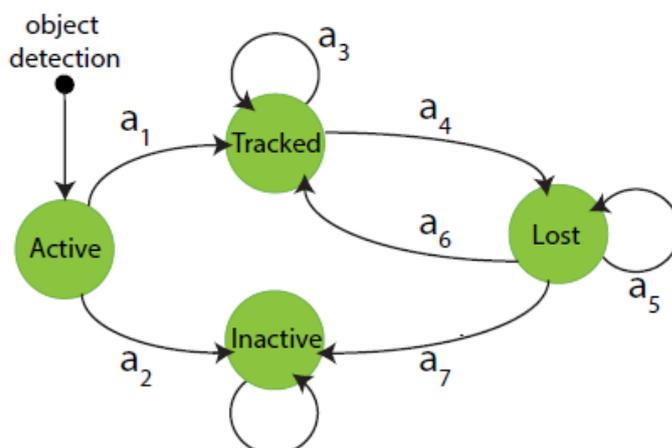
شکل ۲-۸- ردیابی چندعامله براساس الگوریتم مارکوف [۱۷].

فضای حالت MDP در شکل ۲-۹، نشان داده شده است [۱۷، ۱۳].

^۱ window size

^۲Markov Decision Processes

^۳cycle



شکل ۹-۲- فضای حالت الگوریتم مارکوف [۱۷].

باتوجه به شکل ۹-۲، فرایند مارکوف، به چهار زیر بخش Active, Tracked, Lost, Inactive تقسیم می-شوند. که در ادامه شرح داده شده است:

۵-۲. نتیجه گیری

الگوریتم‌های ردیابی براساس مدل و براساس دو بعدی بودن تقسیم‌بندی می‌شوند. در این پایان‌نامه الگوریتم ما براساس مدل، بیشتر شبیه به مدل حرکتی است و براساس ویژگی‌هایی که از اشیاء استخراج می‌شود و با استفاده از جریان نوری، به ردیابی می‌پردازد. همچنین از لحاظ دوبعدی بودن شبیه ردیابی براساس هسته و تحت روش مبتنی بر ویژگی است. در این فصل یک مقایسه از الگوریتم‌های مهم برای تشخیص و ردیابی اشیاء انجام شده است که در روش پیشنهادی ما از الگوریتم تقویتی برای استخراج ویژگی و تشخیص و ردیابی اهداف استفاده شده است.

فصل سوم
روش پیشنهادی

۳-۱. مقدمه

هدف از این پایان‌نامه، ارائه مدلی است که بتوان با استفاده از آن، مشکل رانش در ردیابی شیء را حداقل کرد و این امر مستلزم این است که الگوریتم پیشنهادی در برابر انسداد و نویزهای احتمالی مقاوم باشد و کارایی خود را از دست ندهد. همچنین در فصل‌های قبل ذکر شد که یکی از مشکلات ردیابی زمان واقعی، کمبود حافظه به علت پردازش اطلاعاتیست که در طول زمان ردیابی صورت می‌گیرد. در نتیجه برای حل چنین چالشی تنها چند فریم اخیر، برای ردیابی در نظر گرفته می‌شود و همین امر نیز ممکن است باعث ایجاد مدل نادرست و سبب رانش گردد. الگوریتم پیشنهادی ما، برای کم شدن مشکل رانش راه‌حل‌های ارائه نموده است.

در ادامه فصل، ابتدا ساختار کلی از روش پیشنهادی بیان شده است. سپس به معرفی جزئیات الگوریتم پیشنهادی پرداخته می‌شود. و سپس شرح کاملی از شناسایی اهداف و ردیابی ارائه می‌گردد.

۳-۲. ساختار کلی روش پیشنهادی

در این بخش، به شرح کلی روش پیشنهادی پرداخته می‌شود. برای اینکه در این فشرده‌سازی اطلاعات مفید از دست نرود، چند نمونه از پیش‌زمینه و پس‌زمینه تصویر را انتخاب می‌کنیم که از تلفیق آن‌ها، اطلاعاتی که برای تشخیص و ردیابی شیء مفید است بدست بیاید.

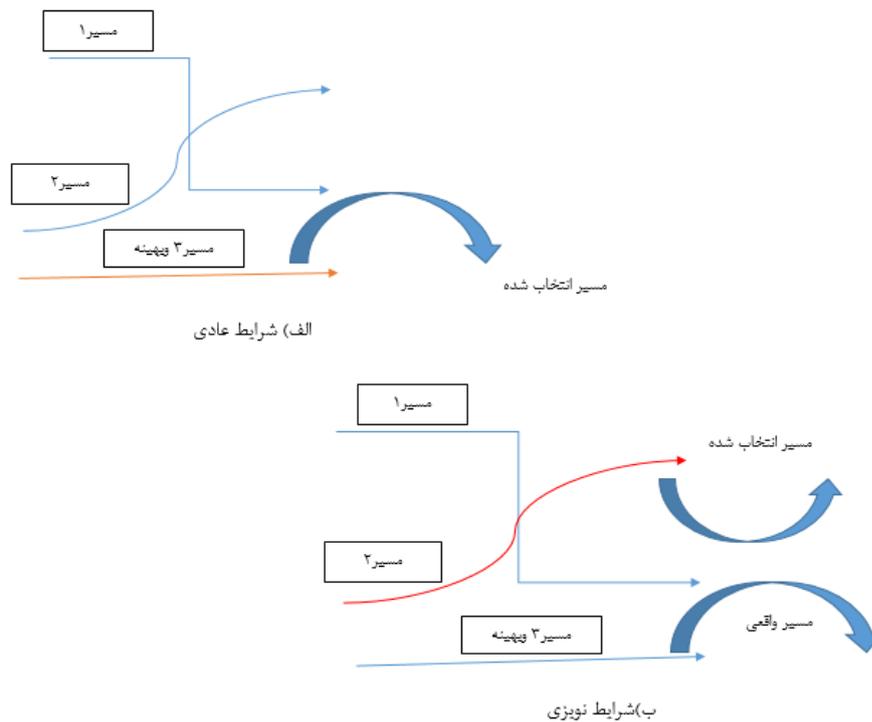
اما برای اینکه ردیابی شیء در مواجهه با نویز و انسدادها دچار رانش نشود باید براساس ویژگی‌ها، یک مدلی از شیء ساخته شود. اما نحوه انتخاب ویژگی یکی از چالش‌های مهم می‌باشد. همچنین در این پایان‌نامه، هدف، ایجاد الگوریتمی هوشمند و پویا می‌باشد که خود بتواند بهترین راه‌حل ممکن را انتخاب کند و انتخاب بهترین راه‌حل ممکن، بستگی به چگونگی انتخاب ویژگی‌ها دارد. درکارهای پیشین، انتخاب ویژگی‌ها هوشمند نبوده است و به صورت خطی، ویژگی‌ها انتخاب می‌شدند. ویژگی‌هایی که به صورت خطی انتخاب می‌شوند یعنی تنها یک مسیر مشخص برای تشخیص و ردیابی شیء دارند، در

مقابل تغییرات ناگهانی مثل انسدادهای شدید، فریم‌ها را گم می‌کنند. در نتیجه الگوریتم دچار رانش می‌شود و نمی‌تواند شیء را به درستی ردیابی کند.

یادگیری تقویتی یک روش غیرنظارت‌شده و در عین حال هوشمند است که به وسیله آزمایش و خطا و پاداش، از میان چندین مسیر(مدل)، مسیر بهتری نسبت به مسیرهای دیگر انتخاب می‌کند و باقی مسیرها هرس می‌شوند. در واقع یادگیری تقویتی یک روش هوشمند غیرخطی است که در مقابل نویزها و انسدادهای شدید از چندین مسیر برای پیدا کردن فریم‌های از دست رفته و ردیابی دوباره هدف، عمل می‌کند. اما یادگیری تقویتی در عین هوشمندی چون تلفیقی از روش حریصانه است، همیشه بهینگی را تضمین نمی‌کند. برای این چالش، در این پایان‌نامه از داده‌های تنک استفاده شده است.

ارتباط بین تنک بودن و بهینگی:

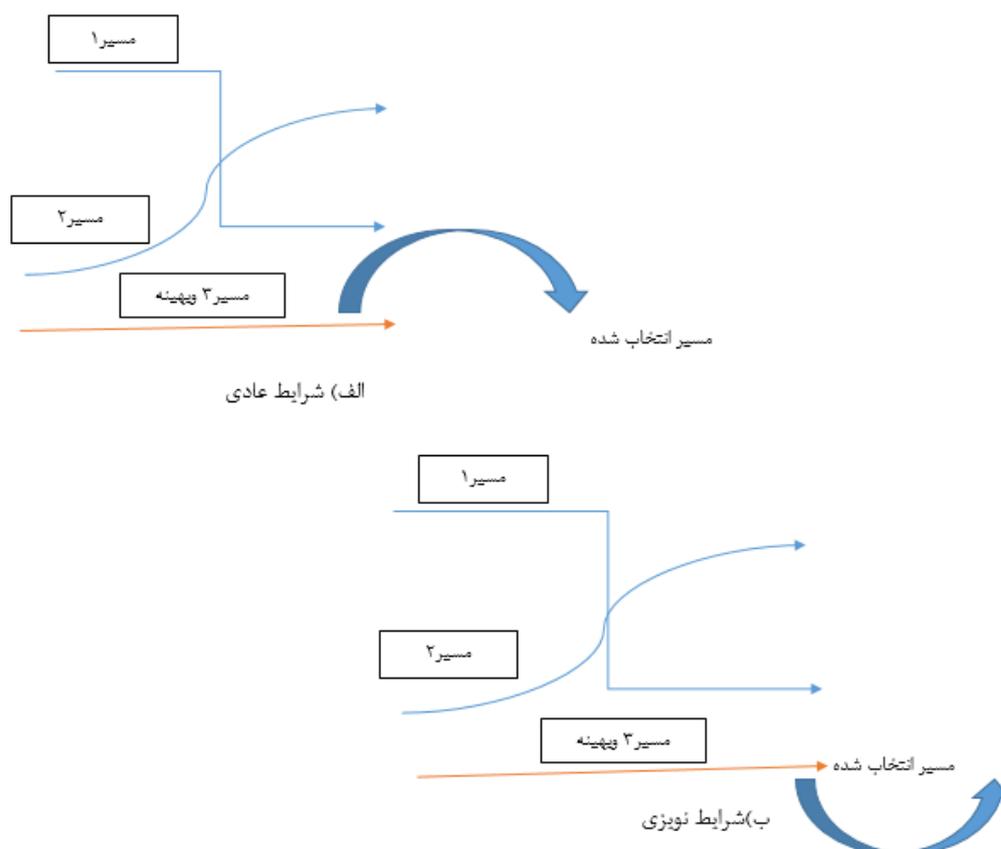
روش پیشنهادی بر این اساس است که یادگیری تقویتی در عین حال که هوشمند است بتواند بهینگی را نیز تضمین کند. الگوریتم‌هایی که وابستگی به داده‌ها دارند در مقابل تغییرات ناگهانی و عواملی مثل نویز و انسداد، کارایی لازم را ندارند. پیش از این در مقاله [۱۷]، الگوریتم تقویتی براساس روش تطبیقی نوشته شده بود. روش‌های تطبیقی، الزاما بهینگی را تضمین نمی‌کنند و ممکن است در شرایط مختلف، دو مدل متفاوت از یک داده ارائه دهند درحالی‌که یک مدلس با وجود نویز، نتواند کارایی لازم در ردیابی داشته باشد اما مدل دیگر در شرایط بدون نویز، کارا باشد. شکل ۳-۳، شماتیکی از مطالب بالا را نشان می‌دهد.



شکل ۳-۱، روش غیرخطی و تطبیقی، در دو شرایط عادی و نویزی

مطابق با شکل ۳-۱، الگوریتم در شرایط عادی و نویزی دو مدل متفاوت از خود ارائه داده است. در شرایط عادی، الگوریتم مسیر ۳ و بهینه را انتخاب می‌کند اما در شرایط نویزی، مسیر ۲، که بهینه نیست انتخاب می‌شود.

در این پایان‌نامه از داده‌های تنک برای ردیابی استفاده شده است. داده‌های تنک، وابسته به داده نیستند و جز روش‌های غیرتطبیقی هستند، در نتیجه در مقابل هر دو شرایط عادی و نویزی، یک مدل یکسان که بهترین مدل ممکن است، از خود ارائه می‌دهند. شکل ۳-۵، شماتیکی از موارد گفته شده بالا را نشان می‌دهد.



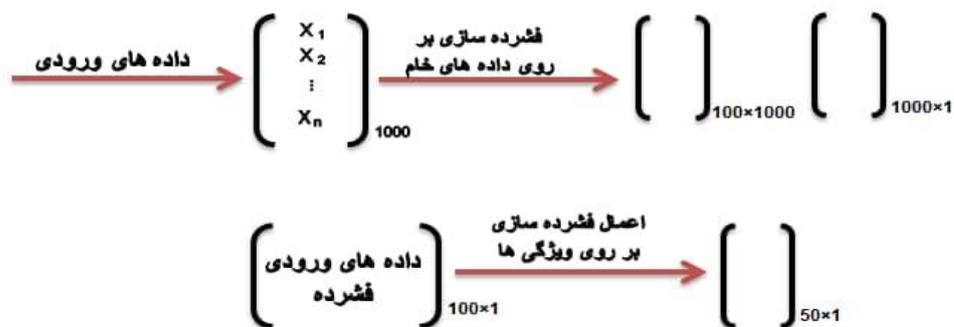
شکل ۲-۳، روش غیرخطی و غیر تطبیقی، در دو شرایط عادی و نویزی

مطابق با شکل ۲-۳، الگوریتم در شرایط عادی و نویزی دو مدل یکسان و کارا، از خود ارائه کرده است.

در هر دو شرایط عادی و نویزی، الگوریتم مسیر ۳ و بهینه را انتخاب می‌کند.

در واقع الگوریتم‌های مبتنی بر تنکی بر اساس درجه تنکی، بهترین ویژگی‌ها را در هر مسیر انتخاب

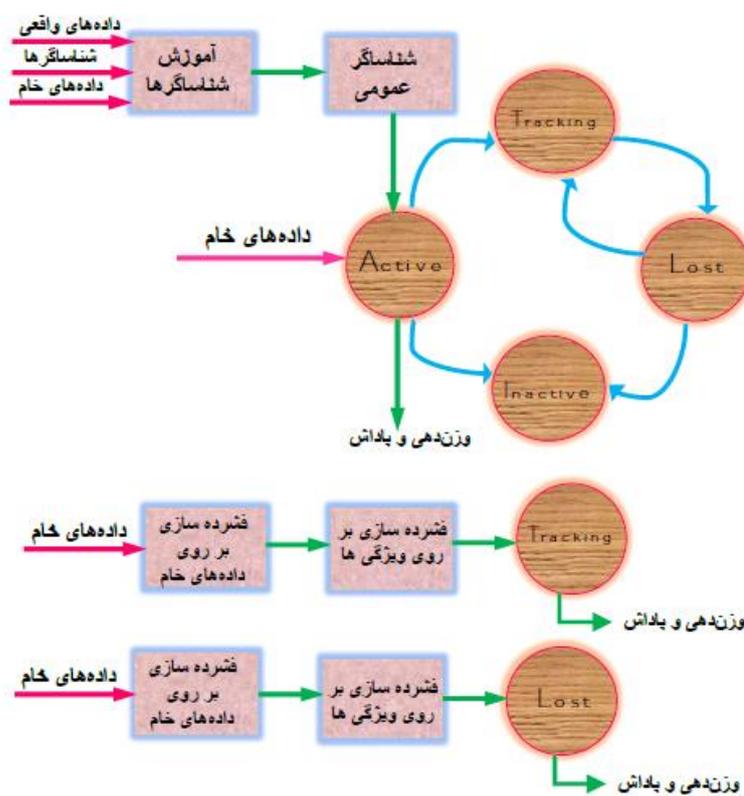
می‌کند تا الگوریتم به علت وجود عواملی چون نویز و انسداد، دچار رانش احتمالی و عدم ردیابی نشود.



شکل ۳-۳ نحوه اعمال فشرده سازی بر روی داده‌های خام و فشرده سازی بر روی ویژگی‌ها

در واقع همان‌طور که در شکل ۳-۳، نشان داده شده است دو مرحله فشرده‌سازی بر روی داده‌ها انجام می‌گیرد.

در این پایان‌نامه برای استخراج ویژگی‌ها از الگوریتم Haar-Like بر مبنای یادگیری تقویتی استفاده شده است. الگوریتم یادگیری تقویتی شامل روش‌های مختلفی است. در اینجا از الگوریتم مارکوف بهره گرفته شده است. الگوریتم مارکوف شامل چهار مرحله برای شناسایی و ردیابی اهداف است که در هر مرحله از یک سیاست خاص برای استخراج ویژگی‌ها استفاده می‌کند. هدف این است که فشرده‌سازی‌ها را در MDP اعمال کنیم. فرایند کلی ردیابی شیء، در شکل ۳-۴ به صورت شماتیک شرح داده شده است.



شکل ۳-۴ نحوه اعمال فشرده‌سازی بر روی ویژگی‌ها در فرایند مارکوف

همان‌طور که در شکل ۳-۴ نشان داده شده است، ابتدا سه ورودی، داده‌های واقعی، شناساگرهای تصویر و داده‌های خام، برای آموزش شناساگرها استفاده می‌شوند. سپس بعد از آموزش شناساگرها، شناساگرهای

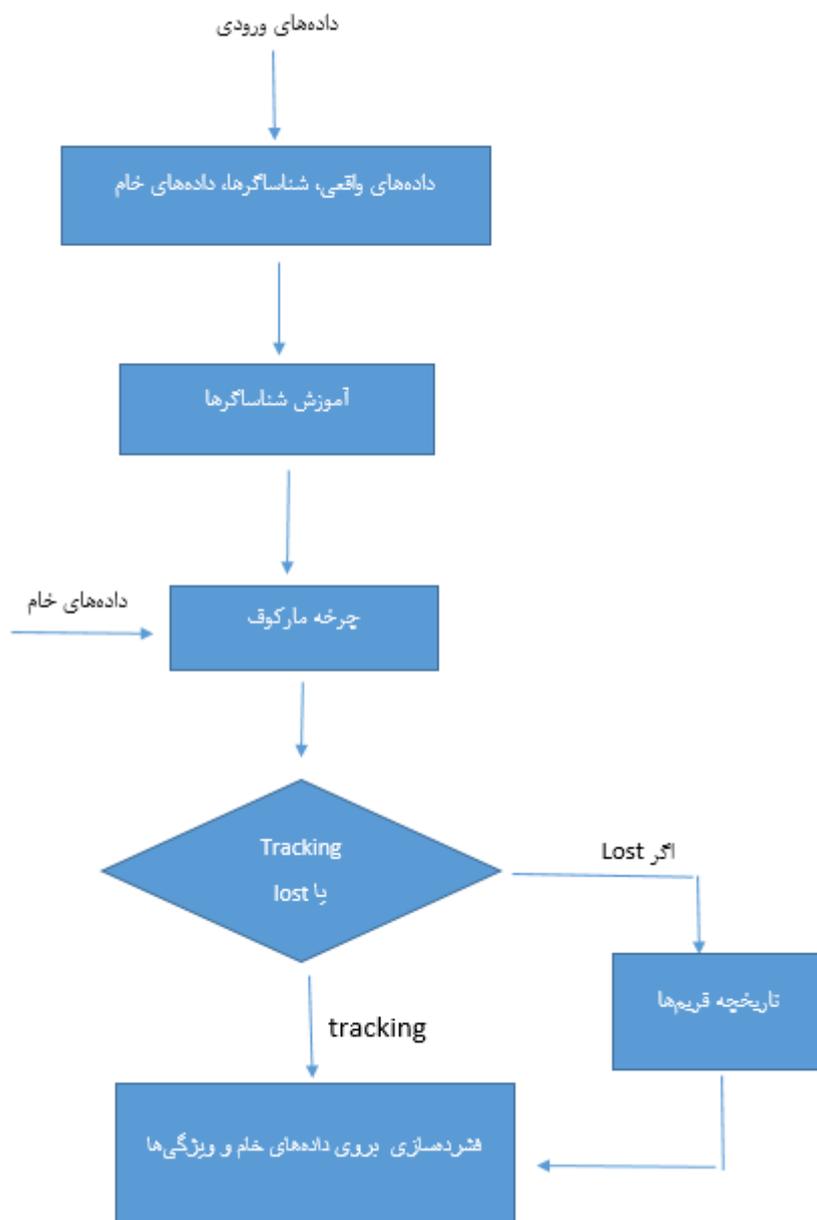
عمومی به همراه با داده‌های خام وارد چرخه‌ی مارکوف می‌شوند. هدف از چرخه مارکوف، ارایه مسیرهای مختلف برای ردیابی شیء و دور زدن نویز یا انسدادهای احتمالی است. در این تحقیق، هدف ردیابی همزمان چندین هدف است. اما الگوریتم مارکوف به صورت تک هدف، اشیاء را شناسایی و ردیابی می‌کند. برای اینکه همزمان چندین شیء ردیابی شوند، الگوریتم ردیابی به صورت پایپلاین، برای هر شیء یک فرایند مارکوف جداگانه در نظر می‌گیرد.

حالت Active ابتدایی‌ترین حالت فرایند مارکوف است که هر شیء‌ی ابتدا وارد این مرحله می‌شود و در اینجا اشیاء، ابتدا مورد شناسایی و آموزش قرار می‌گیرند. و شناساگرهای خاص، از میان شناساگرهای عمومی براساس اشیاء هدف، انتخاب می‌شوند.

در حالت Tracking، داده‌هایی که در مرحله اول (Active) شناسایی شدند مورد ردیابی قرار می‌گیرند. و تا زمانی که عواملی خارجی مثل نویز، باعث برهم زدن روند ردیابی نشود، شیء در این مرحله می‌ماند و تا زمانی که از صحنه خارج نشود، ردیابی می‌شود. در اینجا برای تخمین محل بعدی از شیء، نیاز به استخراج ویژگی است. برای استخراج ویژگی در این مرحله از الگوریتم Haar-like استفاده شده است. در حالت Tracking، ابتدا داده‌های خام، تحت الگوریتم Haar-like و براساس درجه تنکی، استخراج ویژگی و فشرده می‌شوند. سپس دوباره بر روی همین داده‌های تنک، عمل فشرده‌سازی دامنه صورت می‌گیرد. اما هنگامی که داده‌ها به علت انسداد دو یا چند شیء یا به علت نویز، ناپدید شوند وارد حالت Lost می‌شوند.

در حالت Lost تلاش می‌شود که فریم‌های از دست رفته دوباره بازیابی شوند. در این بخش هم از یک سیاست مشخص شده برای آموزش مدل استفاده می‌شود و داده‌ها دوباره استخراج ویژگی می‌شوند و فریم‌های بیشتری به عنوان تاریخچه هدف برای آموزش انتخاب می‌شوند. همچنین در این مرحله هم دو فشرده‌سازی، صورت می‌گیرد. در صورتی که بتوان فریم‌های از دست رفته را بازیابی کرد، الگوریتم دوباره وارد حالت Tracking می‌شود و ردیابی را از سر می‌گیرد. اما اگر داده‌ها بازیابی نشوند و الگوریتم

برای مدت زیادی در حالت Lost بماند وارد حالت Inactive می‌شود. در شکل ۳-۵، روش پیشنهادی، به صورت چارت نشان داده شده است.



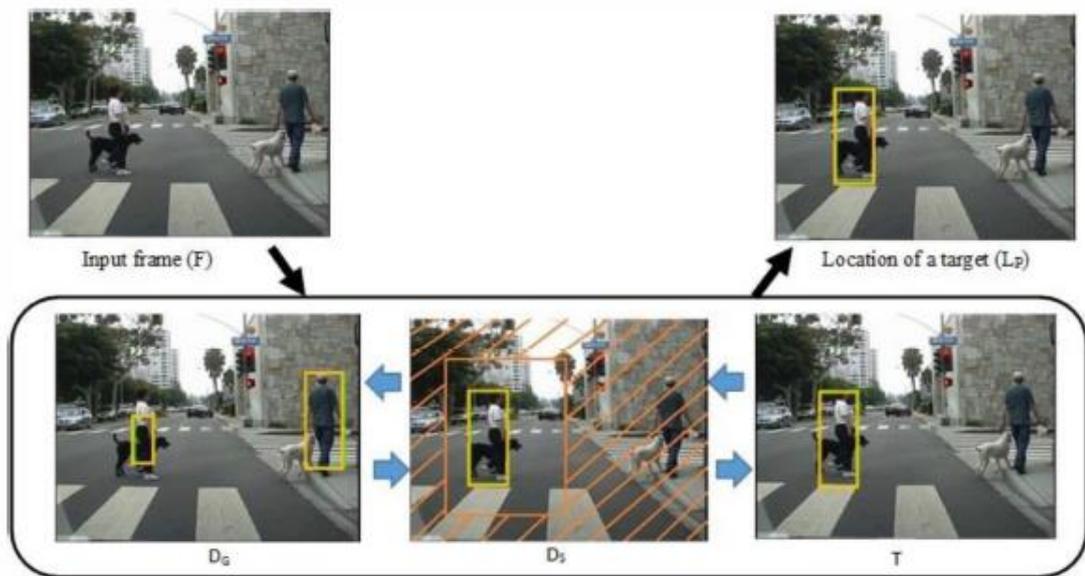
شکل ۳-۵ چارت روش پیشنهادی

۳-۳. شناسایی و ردیابی هدف

پیش از این گفته شده که در این پایان‌نامه برای ردیابی یک شیء، دو مرحله شناسایی و ردیابی نیاز است. در این قسمت، ابتدا در مورد شناساگر هدف و در ادامه در مورد نحوه تخمین مکان بعدی شیء و ردیابی آن شرح داده خواهد شد.

۳-۳-۱. شناساگر عمومی و خاص یک شیء

در اولین مرحله ردیابی یک شیء، بایستی شناسایی صورت گیرد. لذا دو مفهوم شناساگر عمومی و شناساگر خاص مطرح می‌شود. ممکن است در یک صحنه یا یک ویدیو هدف این باشد که مثلاً ۱۰۰ شیء شناسایی و ردیابی شوند. این اشیاء به عنوان شناساگر عمومی و به صورت آفلاین، توسط الگوریتم AdaBoost تحت الگوریتم یادگیری تقویتی، آموزش می‌یابند. سپس در زمان ردیابی هنگامی که یک شیء وارد صحنه می‌شود با شناساگر عمومی که از قبل آموزش داده شده، مطابقت داده می‌شود و اگر با درصد بالایی همپوشانی داشته باشند، آن شیء به عنوان شناساگر خاص، مورد ردیابی قرار می‌گیرد. این مرحله از آموزش نیز توسط الگوریتم AdaBoost انجام می‌گیرد. اگر شناساگر خاص که در زمان t وارد صحنه شده است تا زمان t مورد ردیابی قرار نگیرد، به عنوان ردیاب کاذب شناسایی می‌شود و شیء مورد نظر از شناساگر خاص خارج می‌گردد و اشیاء دیگر به عنوان شناساگر خاص، شناسایی و ردیابی می‌شوند. شکل ۳-۶، یک روش ردیابی به وسیله شناسایی را نشان می‌دهد. که در آن، ابتدا شیء هدف با شناساگرهای عمومی مطابقت داده می‌شود. در صورتی که هدف با یکی از شناساگرهای موجود با درصد بالایی همپوشانی داشته باشد به عنوان شناساگر خاص، مورد ردیابی قرار می‌گیرد.



شکل ۳-۶ روش ردیابی به وسیله شناسایی. شناساگر عمومی (DG)، شناساگر خاص (Ds)، پیش‌بینی موقعیت بعدی یک شناساگر خاص به عنوان هدف [۴].

در واقع اگر، بین شناساگر عمومی و شناساگر خاص (هدف)، درصد همپوشانی 0.5 یا بالاتر باشد، شیء به عنوان هدف جدید، مورد ردیابی قرار می‌گیرد. در رابطه (۳-۱) نشان داده شده است.

$$\text{Overlapping area} = \frac{\text{area}(D_S \cap D_G)}{\text{area}(D_S \cup D_G)} \quad (1-3)$$

در رابطه (۳-۱)، اشتراک شناساگر عمومی (DS) بر روی شناساگر خاص (DG)، تقسیم بر اجتماع فضای تحت پوشش این دو شناساگر، درصد همپوشانی را تعیین می‌کند [۴].

۳-۳-۲. ردیابی هدف

بعد از اینکه شناساگر خاص مورد شناسایی قرار گرفت، ردیابی می‌شود. در این پایان‌نامه از جریان‌نوری استفاده شده که جز روش‌های مدل حرکتی است. و براساس نوع حرکت شیء به ردیابی می‌پردازد. جریان‌نوری، ابتدا چندین نقطه از ویژگی‌های مدل حرکتی یک شیء را در نظر می‌گیرد، سپس محل بعدی قرارگیری شیء را براساس میزان همپوشانی آن نقاط با مکان قبلی می‌سنجد و اگر درصد بالایی همپوشان بودند به عنوان محل بعدی تشخیص داده می‌شود. به عنوان مثال، موقعیت اولیه $U=(u_x, u_y)$

و موقعیت بعدی شیء را V در نظر می‌گیریم. فاصله بین این دو موقعیت را d می‌نامیم، در نتیجه موقعیت بعدی یک شیء از رابطه $V=u+d=(u_x+d_x, u_y+d_y)$ بدست می‌آید. الگوریتم جریان نوری برای اینکه این همپوشانی را انجام دهد از روش forward-backward(FB)error استفاده می‌کند. و اگر موقعیت کنونی شیء را u و موقعیت بعدی آن را u' بنامیم درصد خطای همپوشانی یک شیء به صورت رابطه (۲-۳) بدست می‌آید:

$$e(u) = \| u - u' \|^2 \quad (2-3)$$

برای همه نقاط، طبق رابطه (۳-۳)، همپوشانی محاسبه می‌شود:

$$e_{\text{medFB}} = \text{median}(\{e(u_i)\}_{i=1}^n) \quad (3-3)$$

که در رابطه (۳-۳)، n تعداد نقاط مورد نظر برای همپوشانی هستند. و $e(u)$ درصد همپوشانی به ازای هر یک از نقاط i می‌باشد. اگر درصد خطای مجموع همه این نقاط از درصد تعیین شده کمتر باشد، موقعیت بعدی پایدار نامیده می‌شود. در غیر این صورت ناپایدار است و باید موقعیت دیگری در نظر گرفته شود.



شکل ۳-۷ نمونه‌ای از ردیابی یک شیء به وسیله الگوریتم جریان نوری [۱۷].

در شکل ۳-۷، فریم ۵۰ به عنوان فریم هدف تعیین می‌گردد، سپس فریم ۵۱ به عنوان موقعیت بعدی شیء، همپوشانی می‌شود و چون نقاط به درستی همپوشانی شده است پس پایدار است. اما فریم ۵۷ به درستی همپوشانی نشده است پس ناپایدار بوده و ردیابی دچار رانش شده است. جریان نوری در زمان‌هایی

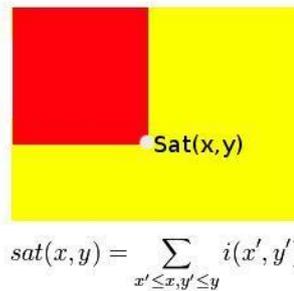
که پس‌زمینه ثابت باشد ممکن است به درستی ردیابی نکند و نقاط ناهمپوشان ایجاد گردد. این امر منجر به گم شدن فریم‌ها می‌شود. در چنین مواقعی علاوه بر در نظر گرفتن نقاط همپوشان، چند فریم قبل را هم به عنوان تاریخچه هدف در نظر می‌گیرد. رابطه (۴-۳)، نقاط همپوشان را با در نظر گرفتن چند فریم از تاریخچه هدف به صورت زیر، محاسبه می‌کند.

$$O_{\text{mean}} = \text{mean}(\{o(t_k, D_k)\}_{k=1}^k) \quad (4-3)$$

k تعداد فریم‌های قبل از فریم کنونی است که به عنوان تاریخچه شیء استفاده می‌شود و D_k شناساگر و t_k به عنوان هدف در زمان k ام می‌باشد.

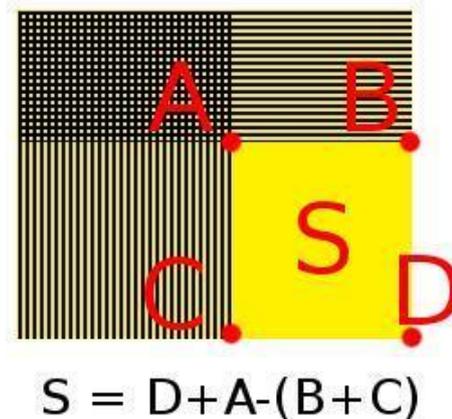
۴-۳. الگوریتم Haar-like

این الگوریتم، یک روش برای استخراج داده‌ها می‌باشد. الگوریتم Haar-like با بکارگیری از تصویر انتگرالی (Integral Image)، این کار را انجام می‌دهد. عملکرد این روش بر این مبناست که تصویر انتگرالی در هر نقطه آن مجموع تمام نقاط ما قبل خود را دارد. شکل ۸-۳، نمایشی از یک تصویر انتگرالی می‌باشد.



شکل ۸-۳ نمایش تصویر انتگرالی

در شکل ۸-۳، مقدار نقطه سفید، در حقیقت مجموع تمام نقاطی است که در مستطیل قرمز قرار گرفته است. حال می‌توان به همین طریق، تصویر انتگرالی سایر نقاط را نیز بدست آورد. شکل ۹-۳، نحوه قرارگیری نقاط در تصویر انتگرالی را نشان می‌دهد.



شکل ۳-۹ نحوه قرارگیری نقاط در تصویر انتگرالی

روش‌های مولفه‌ای وابسته به مولفه‌های مثل رنگ (RGB) و اندازه هستند، در نتیجه باعث می‌شود الگوریتم استخراج ویژگی براساس وابستگی به داده شکل گیرد که این نوع روش‌ها در مقابل نویز و تغییرات ناگهانی تصویر، مقاوم نیستند در نتیجه باعث رانش می‌شوند. اما الگوریتم haar-like براساس ساختار ویژگی‌ها چیده شده است. در نتیجه چون وابسته به داده‌ها نیستند، در مقابل نویز و تغییرات ناگهانی بیشتر مقاوم هستند.

۳-۵. الگوریتم فشرده‌سازی

مبنای اصلی این پایان‌نامه، اعمال فشرده‌سازی بر روی داده‌ها می‌باشد. فشرده‌سازی بر روی داده‌های خام، باعث می‌شود بسیاری از داده‌هایی که مورد نیاز برای ردیابی نیستند، بکاربرده نشود و همچنین این مساله، از پیچیدگی محاسباتی می‌کاهد و باعث بهبود سرعت و کاهش فضای حافظه می‌شود. فشرده‌سازی بر روی داده‌های تنک یا همان فشرده‌سازی بر روی دامنه، باعث می‌شود تنها ویژگی‌های شاخص برای مدل‌سازی برنامه استفاده شود و باقی ویژگی‌ها در نظر گرفته نشود، این امر باعث می‌شود که ردیابی در شرایط مختلف، از جمله تصاویر نویزی و انسداد شیء‌ها، از دقت عمل بالاتری برخوردار باشد و رانش را به حداقل برساند. در این پایان‌نامه هم فشرده‌سازی بر روی داده‌های خام و فشرده‌سازی بر روی ویژگی‌ها، از روش بازسازی تنک و براساس ماتریس اندازه، استفاده شده است.

ماتریس اندازه، با ایجاد داده‌های تصادفی و ضرب در ماتریس داده، فضای حافظه را کاهش می‌دهد در اینجا فرض می‌شود ماتریس $\mathcal{R} \in \mathbb{R}^{n \times m}$ به عنوان ماتریس اندازه در نظر گرفته شود که در آن، n تعداد ویژگی‌ها و m اندازه داده‌ها باشد و اگر $x \in \mathbb{R}^m$ ماتریس داده باشد آنگاه حاصل ضرب فشرده این دو ماتریس از طریق رابطه (۸-۳) بدست می‌آید:

$$v = \mathcal{R}x \quad (8-3)$$

که n بسیار کوچک‌تر از m است. در این فشرده‌گی علیرغم اینکه بعد کاهش می‌یابد، اطلاعات داده‌ای هم از دست نمی‌رود و درصد خطای آن بسیار پایین است. در این پایان نامه، تعداد ویژگی‌های فشرده شده از دامنه، ۵۰ در نظر گرفته شده است. که در فصل ۴، در مورد افزایش یا کاهش این مقدار ویژگی، و تاثیری که بر ردیابی اشیاء می‌گذارد، آزمایش و بحث می‌گردد.

در اینجا، هر دو نوع فشرده‌سازی در الگوریتم Haar-like انجام می‌شود. همچنین الگوریتم Haar-like در فرایند مارکوف در دو بخش Tracking و Lost اعمال می‌شود. هنگامی که از میان شناساگرهای عمومی، شیء هدف برای ردیابی انتخاب می‌شود وارد مرحله Tracking می‌شود. در این مرحله بعد از الزامات اولیه لازم وارد تابع apply_motion_prediction می‌شود. در این تابع، براساس الگوریتم Haar-like و با تلفیق هر دو نوع فشرده‌سازی، موقعیت بعدی شیء، تخمین زده می‌شود. در زیر شمای کلی از نحوه کار تابع apply_motion_prediction نشان داده شده است.



خروجی: تخمین موقعیت بعدی شیء

پارامترهای ثابت:

۱- اندازه ابعاد فشرده شده ویژگی‌ها: $m=50$

۲- ضریب یادگیری: $l=0.89$

۳- اندازه پنجره: $w=20$

مراحل اجرای الگوریتم:

۱- دریافت فریم کنونی و موقعیت اولیه شیء هدف

۲- فشرده‌سازی اولیه بر روی فریم کنونی توسط ماتریس اندازه

۳- محاسبه موقعیت پیکسل‌ها در فریم کنونی توسط تصویر انتگرالی

۴- فشرده‌سازی بر روی ویژگی‌های تنک و وزن‌دهی توسط الگوریتم haar-like

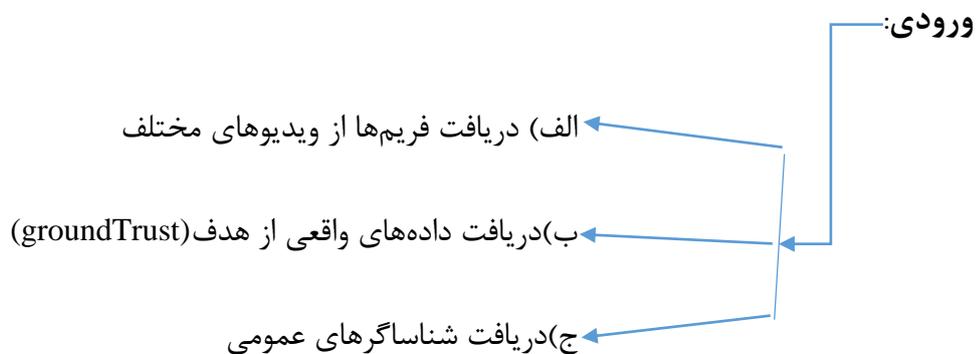
۵- محاسبه طبقه‌بند برای همه نمونه‌ها، در اینجا ده نمونه برای هر شیء در نظر گرفته شده است

۶- پیدا کردن بهترین طبقه‌بند (ماکسیمم)

۷- تخمین موقعیت بعدی شیء هدف

۳-۶. الگوریتم توسعه یافته تقویتی

در زیر شمای کلی از الگوریتم توسعه یافته تقویتی نشان داده شده است:



پارامترهای ثابت:

۱- همپوشانی شیء با شناساگرهای عمومی ۰/۵ در نظر گرفته می‌شود.

۲- برای تخمین محل بعدی شیء، اگر موقعیت شیء کنونی با شیء تخمین زده شده، ۰/۶ درصد یا بالاتر همپوشانی داشته باشد، عمل ردیابی ادامه پیدا می‌کند اگر پایین‌تر از این باشد وارد حالت Lost می‌شود.

۳- اندازه پنجره ۲۰ در نظر گرفته می‌شود.

۴- تعداد ویژگی‌ها برای جریان نوری ۳۰ در نظر گرفته می‌شود.

خروجی: طبقه‌بند باینری (w,b) برای داده‌ها

مرحل اجرای الگوریتم:

۱- وزن‌دهی اولیه w_0 ← w , b_0 ← b

۲- ورود به حلقه

۳- تعیین موقعیت اولیه فریم‌ها

۴- ورود به MDP و رفتن به حالت Active

۵- استخراج ویژگی و تعیین شناساگرهای خاص از میان شناساگرهای عمومی

۶- وزن‌دهی و پاداش براساس سیاست اتخاذشده در حالت Active

۷- رفتن به حالت Tracking

۸- فشردن‌سازی بر روی داده‌های خام و استخراج ویژگی براساس الگوریتم Haar-like

۹- فشردن‌سازی بر روی ویژگی‌ها

۱۰- وزن دهی و پاداش براساس سیاست اتخاذ شده در حالت Tracking

۱۱- ردیابی شیء، در صورت ردیابی کامل، رفتن به گام ۱۵ و در غیر این صورت، رفتن به گام ۱۲

۱۲- حالت Lost

۱۳- آموزش دوباره براساس تاریخچه فریم‌ها و رفتن به گام‌های ۹ و ۸

۱۴- تصمیم‌گیری، برای رفتن به حالت ۷ یا ۱۵

۱۵- ورود به حالت Inactive، پایان ردیابی شیء و رفتن به حالت ۴

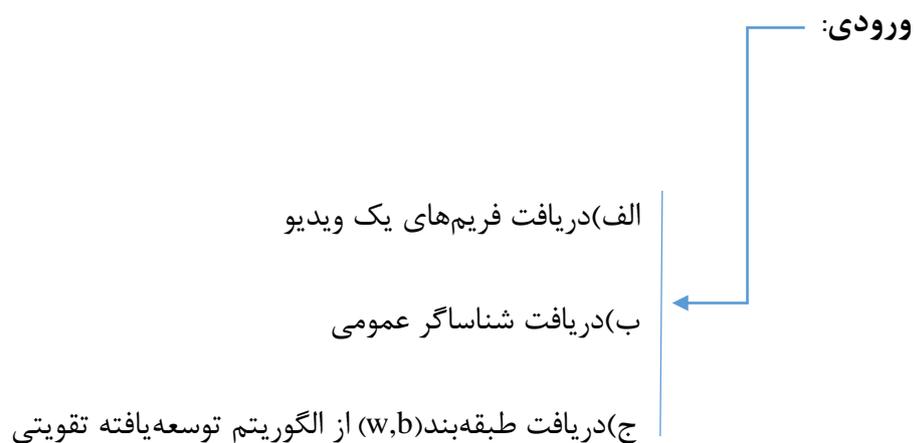
۱۶- پایان حلقه، زمانی که همه‌ی اهداف ردیابی شوند.

۳-۷. الگوریتم کلی ردیابی

این پایان‌نامه براساس ردیابی چندتایی، به ردیابی همزمان چندین شیء می‌پردازد. بعد از اینکه شیء،

شناسایی شد به عنوان یک شناساگر خاص، مورد ردیابی قرار می‌گیرد.

در زیر نحوه کار کلی الگوریتم ردیابی نشان داده شده است:



خروجی: ردیابی همه‌ی اهداف تعیین شده در ویدیو

۱- اهداف ردیابی شده: $T \leftarrow \emptyset$

۲- حالت Tracking در فرایند مارکوف

۳- وزن دهی و پاداش براساس یک سیاست خاص

۴- ردیابی شیء در ویدیو در صورت ناپدید شدن فریم‌ها و رفتن به گام ۵، یا در صورت ردیابی کامل، رفتن به گام ۸ و ۹

۵- حالت Lost در فرایند مارکوف

۶- وزن دهی و پاداش و رسیدن به یک طبقه بند (w,b) براساس تاریخچه هدف

۷- تصمیم‌گیری برای رفتن به گام ۴ یا رفتن به گام ۸

۸- حالت Inactive در فرایند مارکوف و رفتن به گام ۲

۹- $T \leftarrow TU\{t\}$

۱۰- ردیابی همه اهداف و پایان الگوریتم

۳-۸. نتیجه‌گیری

در روش پیشنهادی ما از الگوریتم یادگیری تقویتی که روش غیرخطی است برای ردیابی اشیاء استفاده شده است. و همواره در هر فریم، بسته به شرایط موجود، از میان چندین مدل یکی از مدل‌ها را برای تشخیص و ردیابی اشیاء استفاده می‌کند. یادگیری تقویتی همواره بهترین مسیر را انتخاب نمی‌کند و ممکن است بهینگی را رعایت نکند برای این منظور، از داده‌های تنک استفاده می‌کنیم و با استفاده از الگوریتم haar-like انتخاب ویژگی روی داده‌ها انجام می‌دهیم. این کار باعث می‌شود الگوریتم تقویتی همواره بهینگی را رعایت کند و بهترین مسیر را برای ردیابی و یا بازیابی فریم‌های از دست رفته، استفاده کند. انتخاب داده‌های تنک، خود یک عمل فشرده‌سازی بر روی داده‌های خام، انجام می‌دهد ولی برای اینکه دقت انتخاب ویژگی و مدل ما، بالاتر رود دوباره عمل فشرده‌سازی روی داده‌های تنک، انجام می‌دهیم این کار باعث می‌شود ردیابی از دقت بالاتری برخوردار باشد و رانش کمتری صورت بگیرد.

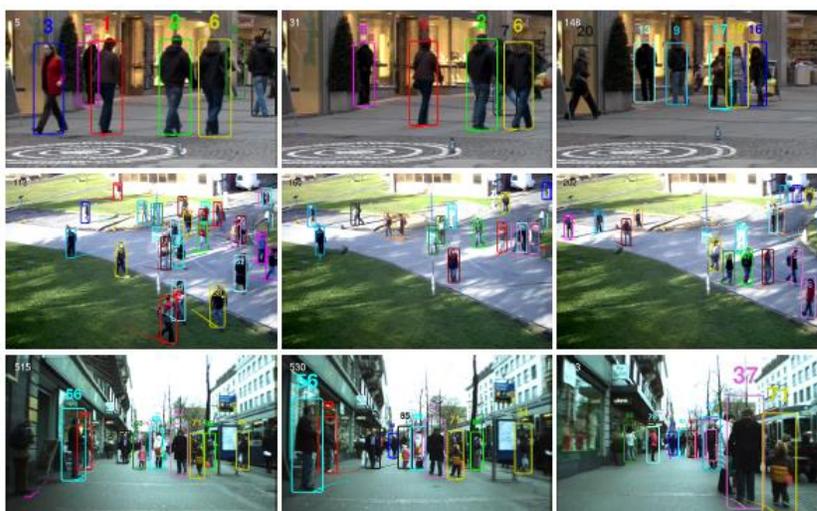
فصل چهارم
نتایج و آزمایش‌ها

۴-۱. مقدمه

در این فصل به بررسی نتایج حاصل از الگوریتم پیشنهادی در شرایط مختلفی از جمله ردیابی چند شیء به طور همزمان در شرایط عادی، نویز و انسداد خواهیم پرداخت. و ویدیوهایی مختلف از پایگاه داده را در این الگوریتم پیاده می‌کنیم و نتایج آن را با الگوریتم‌های پیشین مقایسه خواهیم کرد.

۴-۲. شرایط انجام آزمایشات

پایگاه داده مورد استفاده در این تحقیق شامل ۳۵۰ هزار محدوده جعبه‌ای عابرین برچسب‌گذاری شده در ۲۵۰ هزار فریم از ویدیو با رزولوشن 640×480 می‌باشد. در شکل (۴-۱)، چند نمونه از ویدیوهای پایگاه داده مورد استفاده نشان داده شده است. طول مجموعه داده ۱۳۷ دقیقه مربوط به بخش ویدیو است و اطلاعات از ۱۹۰۰ نفر به دست آمده است. پایگاه داده شامل ۲۲ بخش می‌باشد که ۱۱ بخش مربوط به آموزش و باقی بخش‌ها برای ارزیابی ردیابی چندنفره استفاده می‌شود.



شکل ۴-۱، چند نمونه از ویدیوهای موجود در پایگاه داده مورد ردیابی در پایان نامه

نتایج حاصل از روش پیشنهادی ما، با ۶ الگوریتم ردیابی چندنفره، که در ادامه هر یک به طور مختصر توضیح داده خواهند شد، مقایسه شده است. الگوریتم DP_NMS در [۵۷]، یک الگوریتم حریم‌بندی بهینه سراسری است که با استفاده از برنامه‌نویسی پویا، به صورت خطی به ردیابی اشیاء می‌پردازد. در [۵۸] از الگوریتم TC_ODAL^۱ استفاده شده است که با استفاده از روش متمایزکننده و براساس طبقه‌بندی، اشیاء را ردیابی می‌کند.

الگوریتم SMOT^۲ در [۵۹]، یک الگوریتم کارا برای ردیابی اشیاء می‌باشد. زمانی که فریم‌های هدف بخاطر انسداد یا نویز، از دست می‌رود با بدست آوردن نقاط مشابه در شیء‌ها و با شبیه‌سازی این نقاط، اطلاعات از دست رفته را پیدا می‌کند. در [۶۰] از الگوریتم RMOT^۳ استفاده شده، که الگوریتم کارایی در مقابل تغییرات غیرمنتظره دوربین و بازیابی اطلاعات می‌باشد. الگوریتم CEM^۴ در [۶۱]، یک الگوریتم کارا برای تشخیص و ردیابی شیء و همچنین بهینه‌سازی در مسیر ردیابی است. در اینجا با استفاده از روش کاهش ماندگاری گسسته، سعی می‌شود از طریق نقاط گسسته موجود در مسیر ردیابی، بهینه‌ترین نقاط را برای یکپارچه‌سازی مسیر انتخاب کند. در [۱۷] از روش MDP REL^۵، که یک روش یادگیری تقویتی براساس سیاست مارکوف می‌باشد استفاده شده است.

^۱dynamic programmingNMS

^۲Tracklet confidenceODAL

^۳Similar MultiObject Tracking

^۴Relative Motion Network Multi-object tracking

^۵Continuous Energy Minimization

^۶Markov Decision Processes Real

۳-۴. بررسی صحت تشخیص و ردیابی در مرحله آموزش و آزمایش

در این پایان نامه، ابتدا با داده‌های واقعی و با استفاده از شناساگرهای عمومی، به آموزش اشیاء مورد ردیابی پرداخته شده است. و الگوریتم پیشنهادی ما، توانسته است به درستی مرحله آموزش و آزمایش اشیاء را انجام دهد. شکل (۲-۴) تشخیص و ردیابی اشیاء در مرحله آموزشی را نشان می‌دهد.



شکل ۲-۴ تشخیص و ردیابی اشیاء در مرحله آموزشی

شکل (۲-۴)، قسمت الف)، داده‌های واقعی از نتایج هستند. یعنی زمانی که از الگوریتم پیشنهادی استفاده می‌کنیم در واقع یک شبیه‌سازی صورت می‌گیرد که با استفاده از این شبیه‌ساز می‌خواهیم داده‌ها را ردیابی

کنیم هرچقدر شبیه‌ساز ما از دقت بالاتری برخوردار باشد باید به داده‌های واقعی نزدیکتر باشد. نتایج شبیه‌سازی در قسمت ج) نشان داده شده است و نزدیکی نتایج شبیه‌سازی شده به داده‌های واقعی از نتایج، کارایی الگوریتم را نشان می‌دهد. در قسمت ب)، شناساگرهای عمومی می‌باشند. ابتدا شناساگرهای عمومی آموزش داده شده و سپس شناساگرهای خاص برای ردیابی انتخاب می‌شوند. در هر مرحله یک شناساگر خاص انتخاب و آموزش داده می‌شود. در قسمت ج) شناساگر خاص انتخاب شده و عمل ردیابی صورت می‌گیرد. در قسمت د) اگر عواملی چون نویز و انسداد، موقتا باعث عدم ردیابی شیء شود، شناساگر بعدی موقتا برای ردیابی انتخاب می‌شود و شیء گم شده وارد حالت Lost می‌شود و سعی در بازیابی فریم‌های از دست رفته می‌کند، که در شکل (۳-۴) نشان داده شده است. قسمت الف) داده‌های واقعی را نشان می‌دهد و در قسمت ب) شناساگرهای عمومی که برای هر فریم شناسایی می‌شود. اما همان‌طور که در قسمت ج) نشان داده شده، ردیابی به علت نویز یا انسداد دچار مشکل شده و بخشی از فریم‌ها از دست رفته است. برای این منظور شناساگر خاص، موقتا به حالت Lost در قسمت د) می‌رود و عمل بازیابی بر روی آن صورت می‌گیرد و زمانی که بازیابی با موفقیت انجام شود دوباره شیء کنونی ردیابی را از سر می‌گیرد. این روند آموزشی بارها و بارها تکرار می‌شود تا زمانی که همه اشیاء مورد ردیابی به درستی شناسایی و ردیابی شوند.



شکل ۴-۳ تشخیص و ردیابی اشیاء در مرحله آموزشی، در این مرحله شیء وارد حالت lost شده است.

در مرحله آزمایش، داده‌های واقعی از نتایج وجود ندارند و باید شیء ما با داده‌های خام ورودی، آموزش یابند و اشیاء را شناسایی و ردیابی کنند. همچنین در این مرحله همزمان همه اهداف باهم ردیابی می‌شوند. که در شکل (۴-۴)، نشان داده شده است. همان‌طور که در این شکل می‌بینیم، در قسمت الف) ابتدا در هر فریم، شناساگرهای عمومی شناسایی می‌شوند سپس در قسمت ب) شیءها بطور همزمان و فریم به فریم ردیابی می‌شوند. و در قسمت ج) اگر فریمی به علت نویز و انسداد یا زمانی که از ویدیو خارج شود وارد این مرحله می‌شود.



شکل ۴-۴ تشخیص و ردیابی اشیاء در مرحله آزمایش

۴-۴. بررسی تغییرات شیء از لحاظ معیارهای FP و FN

معیارهای منفی کاذب (FN) و منفی حقیقی (FP)، به عنوان دو معیار مهم، برای یک سیستم ردیاب، برای تشخیص و ردیابی شیء می باشد. اما قبل از بررسی نتایج حاصل از این دو معیار، به تعریف آن‌ها می پردازیم.

۴-۴-۱. مقایسه با مقاله پایه^۳

در این قسمت به بررسی دو معیار منفی کاذب و منفی حقیقی می پردازیم. هرچقدر این دو معیار، کمتر باشند، الگوریتم کارایی بالاتری دارد. در این بخش، ابتدا از چهار ویدیو از پایگاه داده که نسبت به بقیه دارای

^۱False positives

^۲False negatives

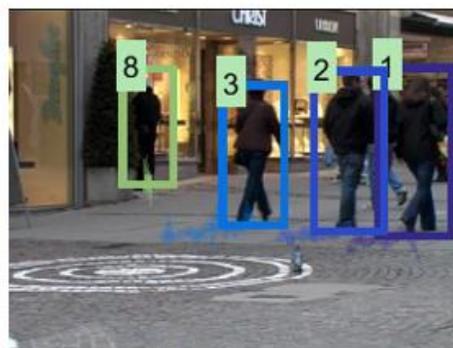
^۳[۱۷]

^۴TUD-Stadtmitte, TUD-Campus, ETH-Sunnyday, ETH-Pedcross2

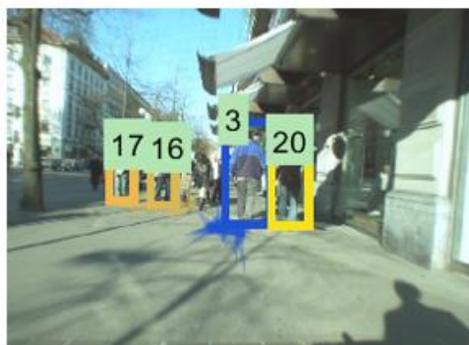
نویز و انسداد بیشتری هستند برای هر دو بخش آموزش و آزمایش انتخاب می‌شوند. در شکل (۴-۵)، برخی از فریم‌های ویدیوها نمایش داده شده است.



فریم ۵۳ از ویدیو TUD-Campus



فریم ۴۱ از ویدیو TUD-Stadtmitte



فریم 298 از ویدیو ETH-Sunnyday



فریم ۴۸۰ از ویدیو ETH-Pedcross2

شکل (۴-۵)، نمایش برخی از فریم‌های مورد ردیابی

بطور جداگانه، بر روی هر یک از ویدیوها، عمل آموزش و آزمایش صورت می‌گیرد و سپس نتایج حاصل از الگوریتم پیشنهادی، با نتایج مقاله پایه مقایسه می‌شود. نتایج، در جدول‌های (۴-۱) و (۴-۲) نشان داده شده است.

جدول ۱-۴ نتایج حاصل از FN و FP بر روی ویدیوهایی با انسداد شدید، براساس مقاله پایه

ویدیوها	FP	FN
TUD-Stadtmitte	۸۴	۳۵۷
TUD-Campus	۴۵	۱۲۳
ETH-Sunnyday	۱۰۸	۱۱۶۹
ETH-Pedcross2	۲۷۷	۵۰۸۲

جدول ۲-۴ نتایج حاصل از FN و FP بر روی ویدیوهایی با انسداد شدید، براساس روش پیشنهادی ما، اعدادی که **Bold** در

آمدند نشان دهنده بهتر شدن الگوریتم است.

ویدیوها	FP	FN
TUD-Stadtmitte	۷۱	۳۴۹
TUD-Campus	۳۵	۱۱۷
ETH-Sunnyday	۱۱۱	۷۳۵
ETH-Pedcross2	۲۵۱	۴۰۳۱

همان طور که از جدول های بالا، مشخص است، روش پیشنهادی ما توانسته عملکرد خوبی نسبت به الگوریتم مرجع داشته باشد. ویدیوها از بالا به پایین در جدول از لحاظ انسداد و اهداف مورد ردیابی، بیشتر و سخت تر می شوند. که نتایج آن در جدول بالا، آمده است و عددی که **Bold** درآمدند نشان دهنده بهبود نتایج است.

۴-۴-۳. مقایسه با الگوریتم‌های پیشین

در این قسمت، از ۱۱ ویدیو برای آموزش^۱ و ۱۱ ویدیو از پایگاه داده هم، برای آزمایش^۲ استفاده شده است. که در این بخش ویدیوهای آموزش و آزمایش، متفاوت از یکدیگر هستند و همه ویدیوها به صورت همزمان آموزش و آزمایش روی آن‌ها صورت می‌گیرد. تعداد کل دو معیار منفی کاذب و منفی حقیقی را از میان این ویدیوها بر روی روش پیشنهادی، بدست می‌آوریم و نتایج حاصل را با الگوریتم‌های پیشین مقایسه می‌کنیم. البته نتایج حاصل از الگوریتم‌های پیشین از مقاله [۱۷] گرفته شده است که در جدول (۴-۳)، نشان داده شده است.

جدول ۴-۳ مقایسه نتایج حاصل از دو معیار منفی کاذب و منفی حقیقی با سایر روش‌های پیشین

Tracker	FP	FN
DP_NMS	۱۳۱۷۱	۳۴۸۱۴
TC_ODAL	۱۲۹۷۰	۳۸۵۳۸
SMOT	۸۷۸۰	۴۰۳۱۰
RMOT	۱۲۴۷۳	۳۶۸۳۵
CEM	۱۴۱۸۰	۳۴۵۹۱
MDP_REL	۹۷۱۷	۳۲۴۲۲
MDP_SPARSE(ours)	۷۵۴۱	۲۹۰۲۸

^۱TUD-Stadtmitte, TUD-Campus, PETS09-S2L1, ETH-Bahnhof, ETH-Sunnyday, ETH-Pedcross2, ADL-Rundle-6, ADL-Rundle-8, KITTI-13, KITTI17, Venice-2

^۲TUD-Crossing, PETS09-S2L2, ETH-Jelmoli, ETHLinthescher, ETHCrossing, AVG-TownCentre, ADL-Rundle-1, ADL-Rundle-3, KITTI-16, KITTI-19, Venice-1

مطابق جدول (۳-۴)، الگوریتم پیشنهادی ما، توانسته است در مجموع، بهبود قابل توجه‌ای نسبت به سایر الگوریتم‌های پیشین در شرایط یکسان داشته باشد.

۴-۴-۴. درصد معیار منفی حقیقی و منفی کاذب

در بالا نحوه بدست آوردن نرخ معیار منفی حقیقی و منفی کاذب، گفته شد. در جدول (۴-۴)، این دو معیار بر روی روش پیشنهادی ما، برای ۱۱ ویدیو آموزش و ۱۱ ویدیو آزمایش بدست آمده است.

جدول ۴-۴ نرخ معیار منفی حقیقی و منفی کاذب در روش پیشنهادی

Tracker	نرخ FP	نرخ FN
MDP_SPARSE(ours)	۲۵/۸ درصد	۱۸/۲ درصد

همان‌طور که نشان داده شده است این دو معیار در الگوریتم پیشنهادی ما، درصد پایینی دارند و هرچقدر در الگوریتمی این درصد پایین‌تر باشد، نشان دهنده کارایی الگوریتم مربوطه است.

۴-۵. بررسی دقت و صحت اشیاء مورد ردیابی

یکی از معیارهای سنجشی مهم، برای یک سیستم ردیاب، بررسی دقت و صحت شیء‌های مورد ردیابی است. اما قبل از بررسی نتایج حاصله، به تعریف این دقت و صحت در ردیابی، می‌پردازیم.

۴-۵-۱. تعریف معیار دقت و صحت در ردیابی اشیاء

صحت (Accuracy): میزان نزدیک بودن نتیجه اندازه‌گیری انجام شده، توسط یک وسیله اندازه‌گیری خاص به اندازه واقعی جسم اندازه‌گیری شده را صحت اندازه‌گیری می‌گویند.

صحت ردیابی چند هدفه (MOTA): صحت ردیابی از رابطه (۴-۱) بدست می آید.

$$MOTA = 1 - \frac{\sum_t (fn_t + fp_t + Idsw_t)}{\sum_t g_t} \quad (1 - 4)$$

در این رابطه، متغیر t فریم کنونی و g_t تعداد اشیاء موجود در آن فریم می باشد. و شامل سه خطا، fn (منفی کاذب) و fp (مثبت کاذب) و $Idsw$ (تغییر هویت) می باشد. تغییر هویت، یعنی مکان هایی از ردیابی که فریم های شیء از دست رفته و ردیاب مجبور شده که شناساگر جدیدی را برای ردیابی پیشنهاد کند.

دقت (Precision): انطباق یا نزدیکی مقادیر اندازه گیری شده به هم را دقت اندازه گیری می گویند. به عبارت دیگر دقت، مقدار تراکنش در اطراف میانگین می باشد.

دقت ردیابی چند هدفه (MOTP): دقت ردیابی از رابطه (۴-۲) بدست می آید.

$$MOTP = \frac{\sum_t d_t^i}{\sum_t c_t} \quad (2 - 4)$$

در این رابطه، متغیر d_t^i فاصله بین مسیر واقعی و مسیر پیش بینی شده شیء i در فریم t و c_t تعداد مسیرهای موجود در فریم t می باشد.

۴-۵-۲. مقایسه با مقاله پایه

در این پایان نامه، همه ی ویدیوهای موجود، مورد آموزش و آزمایش قرار می گیرند. اما هدف و مبنای پایان نامه، کاهش رانش و بهبود دقت در ویدیوهایی که از لحاظ انسداد شیء ها بیشتر می باشند، است. در نتیجه در این قسمت، از میان ویدیوهای موجود، چهار ویدیو که انسداد شیء ها در آنها بیشتر است انتخاب می کنیم.

¹Multiple object tracking accuracy

²Multiple object tracking precision

³TUD-Stadtmitte, TUD-Campus, ETH-Sunnyday, ETH-Pedcross2

بطور جداگانه، بر روی هر یک از ویدیوها، عمل آموزش و آزمایش صورت می‌گیرد و نتایج دقت و صحت ردیابی اشیاء در این ویدیوها را با مقاله پایه [۱۷]، مقایسه می‌کنیم. همچنین برای اطمینان از نتایج مقاله پایه، بر روی این مقاله هم، ویدیوها دوباره آموزش و آزمایش، شده است. نتایج حاصل در جدول (۴-۳) و جدول (۴-۴) نشان داده شده است.

جدول ۴-۵ نتایج حاصل از MOTA و MOTP بر روی ویدیوهایی با انسداد شدید، براساس مقاله پایه

ویدیوها	MOTA (درصد)	MOTP (درصد)
TUD-Stadtmitte	۶۱/۴	۶۶/۵
TUD-Campus	۵۱	۷۱/۲
ETH-Sunnyday	۳۱/۲	۷۶
ETH-Pedcross2	۱۴/۲	۷۱/۳

جدول ۴-۶ نتایج حاصل از MOTA و MOTP بر روی ویدیوهایی با انسداد شدید، براساس روش پیشنهادی ما

ویدیوها	MOTA (درصد)	MOTP (درصد)
TUD-Stadtmitte	۶۲/۲	۶۶/۵
TUD-Campus	۵۴/۳	۷۱/۸
ETH-Sunnyday	۴۷	۷۶/۷
ETH-Pedcross2	۲۶/۱	۷۱/۶

جدول‌های (۴-۴) و (۴-۵)، نتایج حاصل از دقت و صحت ردیابی بر روی چهار ویدیو موجود می‌باشد.

۴-۵-۳. مقایسه با الگوریتم‌های پیشین

در این قسمت، از ۱۱ ویدیو برای آموزش و ۱۱ ویدیو از پایگاه داده هم، برای آزمایش، استفاده شده است. که در این بخش ویدیوهای آموزش و آزمایش، متفاوت از یکدیگر هستند و همه ویدیوها به صورت همزمان آموزش و آزمایش روی آن‌ها صورت می‌گیرد و میانگین دو معیار دقت (MOTP) و صحت (MOTA) را از میان این ویدیوها بر روی روش پیشنهادی، بدست می‌آوریم و نتایج حاصل را با الگوریتم‌های پیشین مقایسه می‌کنیم.

جدول ۴-۷ مقایسه نتایج حاصل از دو معیار دقت (MOTP) و صحت (MOTA) الگوریتم پیشنهادی با سایر روش‌های

Tracker	MOTA(درصد)	MOTP(درصد)
DP_NMS	۱۴/۵	۷۰/۸
TC_ODAL	۱۵/۱	۷۰/۵
SMOT	۱۸/۲	۷۱/۲
RMOT	۱۸/۶	۶۹/۶
CEM	۱۹/۳	۷۰/۷
MDP_REL	۳۰/۳	۷۱/۳
MDP_SPARSE(ours)	۳۶/۷	۷۱/۵

همان‌طور که در جدول (۴-۶)، نشان داده شده، مقدار هر دو معیار صحت و دقت ردیابی، بهبود یافته است.

۴-۶. بررسی MOTA با ویدیوهای تکی

در این قسمت از ویدیوهای آموزشی، ۵ ویدیو برای آموزش و ۶ ویدیو برای آزمایش، استفاده شده است.

جدول (۴-۸) نتایج حاصل از MOTA الگوریتم مقاله [۱۷]

Training Sequences	TUD-Stadtmitte	۵۶.۰	۴۴.۸	۱۴.۰	۲۰.۰	۳۰.۸	۶۰.۸
	ADL-Rundle-6	۴۴.۸	۴۳.۴	۱۳.۳	۲۲.۶	۳۰.۸	۶۰.۳
	ETH-Sunnyday	۴۷.۹	۴۸.۲	۱۱.۵	۲۶.۱	۲۹.۸	۵۷.۸
	KITTI-13	۵۳.۲	۴۷.۵	۱۳.۹	۲۰.۹	۳۲.۱	۵۹.۹
	PETS09-S2L1	۴۹.۰	۴۲.۱	۱۱.۵	۲۲.۱	۲۹.۴	۶۱.۲
		ETH-Sunnyday	ADL-Rundle-8	Venice-2	KITTI-17		
		TUD-Campus	ETH-Pedcross2				
		Testing Sequences					

جدول (۴-۹) نتایج حاصل از MOTA براساس الگوریتم پیشنهادی

Training Sequences	ETH-Sunnyday	۵۹.۰	۵۰.۱	۲۵.۰	۲۶.۶	۳۲.۳	۶۳.۲
	TUD-Stadtmitte	۴۳.۳	۴۰.۲	۲۰	۲۸.۲	۳۱.۷	۶۲.۶
	ADL-Rundle-6	۴۹	۵۱.۱	۱۸.۴	۳۰.۳	۳۳.۹	۵۹.۲
	KITTI-13	۵۱.۵	۴۹.۳	۱۹.۸	۲۷	۳۰.۱	۶۴
	PETS09-S2L1	۵۱.۳	۴۷	۱۸.۶	۲۸.۱	۳۲.۵	۶۱.۹
		ETH-Sunnyday	ADL-Rundle-8	Venice-2	KITTI-17		
		TUD-Campus	ETH-Pedcross2				
		Testing Sequences					

همان طور که از جدول های بالا مشخص است، الگوریتم پیشنهادی در این پایان نامه نسبت به الگوریتم مرجع، بهبود یافته است و عددهایی که به شکل Bold هستند.

۴-۷. بررسی زمان اجرای الگوریتم پیشنهادی

در این بخش، به بررسی زمان اجرای الگوریتم پیشنهادی در هر دو مرحله آموزش و آزمایش می پردازیم. در اینجا از ۱۱ ویدیو برای آموزش و ۱۱ ویدیو از پایگاه داده هم، برای آزمایش، استفاده شده است. که نتایج زمان اجرای آن در جدول (۴-۹)، نشان داده شده است.

جدول (۴-۱۰) نتایج زمان اجرای الگوریتم پیشنهادی در مرحله آموزش و آزمایش

مرحله آزمایش	مرحله آموزش	زمان اجرا
دو ساعت و پنج دقیقه	۵ ساعت و ربع	MDP_SPARSE(ours)

در جدول بالا، زمان اجرای آموزش، تقریباً بیش از دو برابر زمان آزمایش شده است، به این علت است که مرحله آموزش برای اشیاء در ویدیوها، بارها و بارها تکرار می شود تا همه اشیاء به درستی شناسایی و ردیابی شوند.

۴-۸. بررسی الگوریتم در صورت عدم فشردگی سازی در یکی از مراحل Tracking و Lost

در این بخش، به بررسی عدم اعمال فشردگی سازی در یکی از بخش های Tracking یا Lost می پردازیم. چهار ویدیو از پایگاه داده که در بخش های ابتدایی استفاده شدند، مورد بررسی قرار می گیرند. در اینجا دقت و صحت ردیابی مورد آزمایش قرار می گیرد. جدول های (۴-۹) و (۴-۱۰)، نتایج حاصل از آزمایش را در حالت Tracking و Lost نشان می دهد.

جدول ۴-۱۱ نتایج حاصل از MOTP و MOTA در صورت عدم فشرده‌سازی در حالت Tracking

ویدیوها	MOTA (درصد)	MOTP (درصد)
TUD-Stadtmitte	۶۱	۶۶/۵
TUD-Campus	۵۲/۲	۷۱/۳
ETH-Sunnyday	۳۳	۷۵/۱
ETH-Pedcross2	۱۵,۳	۷۱/۴

جدول ۴-۱۲ نتایج حاصل از MOTP و MOTA در صورت عدم فشرده‌سازی در حالت Lost

ویدیوها	MOTA (درصد)	MOTP (درصد)
TUD-Stadtmitte	۵۹	۶۱
TUD-Campus	۴۳/۸	۶۷/۲
ETH-Sunnyday	۲۹	۷۳
ETH-Pedcross2	۱۳/۹	۷۰/۳

ابتدا به بررسی دو معیار دقت و صحت ردیابی در شرایطی که فشرده‌سازی به بخش Tracking اعمال نمی‌شود می‌پردازیم. همان‌طور که از در جدول (۴-۹) نشان داده شده، در این حالت MOTA به نسبت به نتایج حاصل از جدول (۴-۵)، کاهش یافته و نتایج ضعیف‌تری گرفته شده است اما به نسبت به جدول (۴-۴) که فشرده‌سازی در هیچ کدام از بخش‌ها اعمال نمی‌شود درصد بهتری اخذ نموده است. این نشان می‌دهد که فشرده‌سازی تنها در بخش Tracking باعث بهبود نتایج می‌شود. معیار MOTP هم همانند معیار MOTA به نسبت به جدول (۴-۵) که مربوط حالتی است که فشرده‌سازی در هر دو بخش اعمال شده

است، نتایج ضعیف‌تری گرفته است اما به نسبت به جدول (۴-۴) نتایج بهتری اخذ کرده است البته تنها در ویدیو ETH-Sunnyday نتیجه ضعیف‌تری کسب کرده است.

در بالا در مورد نتایج حاصل از عدم فشردگی‌سازی در بخش Tracking صحبت شد. در این قسمت به نتایج حاصل از جدول (۴-۱۰) می‌پردازیم. این جدول حالتی را نمایش می‌دهد که فشردگی‌سازی تنها در بخش Tracking اعمال می‌شود. و طبق پیش‌بینی ما و همچنین طبق نتایجی که بدست آمد، گویای این است که در این حالت صحت و دقت ردیابی حتی نسبت به حالت اولیه در جدول (۴-۴) که هیچ فشردگی‌سازی صورت نمی‌گیرد، نتایج ضعیف‌تری کسب کرده است. علت مساله این است هنگامی که داده‌ها فشردگی‌تری شوند و داده‌های تنک برای فشردگی‌سازی استفاده می‌شوند، مدل‌های دقیق‌تری از ردیابی را ارائه می‌دهند. حال اگر ردیابی دچار مشکل شود و فریم‌ها به علت نویز و انسداد شیء‌ها از دست رود، ردیابی به حالت Lost می‌رود. در این حالت باید مدل ما، مدل همسان یا قوی‌تری نسبت به حالت Tracking باشد تا بتواند با در نظر گرفتن تاریخچه فریم‌ها، فریم‌های از دست رفته را، بازیابی کند. اما در اینجا، در حالت lost برپایه مقاله [17]، مدل‌سازی شده است که مدل ضعیف‌تری نسبت به داده‌های تنک است در نتیجه از کارایی لازم در بازیابی فریم‌های از دست رفته برخوردار نیست برای همین مدل ضعیف‌تری حتی نسبت به حالت اولیه دارد. این نشان از اهمیت حالت Lost می‌دهد که باید مدل قوی‌تر یا حداقل همسان با حالت Tracking داشته باشد وگرنه الگوریتم کارایی خود را از دست می‌دهد.

۴-۹. تاثیر فشردگی‌سازی‌های مختلف بر عملکرد الگوریتم

در این پایان‌نامه ابعاد فشردگی‌شده ویژگی‌ها را، ۵۰ در نظر گرفته شده و براساس آن، جدول (۴-۵) شکل گرفته است. در اینجا دوباره بر روی همان ویدیوها از جدول (۴-۵)، به ازای فشردگی‌سازی‌های مختلف عملکرد الگوریتم را محاسبه می‌کنیم. جدول‌های (۴-۱۱) و (۴-۱۲) نتایج حاصل از دقت و صحت ردیابی از فشردگی‌سازی‌های مختلف را نشان می‌دهد.

جدول ۴-۱۳ نتایج حاصل از MOTA و MOTP با تعداد ویژگی‌های ۴۰

ویدیوها	MOTA (درصد)	MOTP (درصد)
TUD-Stadtmitte	۶۱/۸	۶۶/۲
TUD-Campus	۵۳/۲	۷۱/۸
ETH-Sunnyday	۳۶	۷۶/۳
ETH-Pedcross2	۱۸/۵	۷۰

جدول ۴-۱۴ نتایج حاصل از MOTA و MOTP با تعداد ویژگی‌های ۶۰

ویدیوها	MOTA (درصد)	MOTP (درصد)
TUD-Stadtmitte	۶۱/۲	۶۶/۵
TUD-Campus	۵۱	۷۱/۴
ETH-Sunnyday	۳۹	۷۵/۹
ETH-Pedcross2	۱۷/۷	۷۱/۱

ابتدا فشرده‌سازی ویژگی‌ها را به ازای مقدار ۴۰، با توجه به جدول (۴-۱۱) مورد بررسی قرار می‌دهیم. همان‌طور که از جدول پیداست، نتایج حاصل نسبت به جدول (۴-۵)، در هر دو حالت صحت و دقت ردیابی، عملکرد ضعیف‌تری داشته است.

فشرده‌سازی را این بار به ازای مقدار ۶۰، با توجه به جدول (۴-۱۲) مورد بررسی قرار می‌دهیم. همان‌طور که از جدول پیداست، نتایج حاصل نسبت به جدول (۴-۵)، در هر دو حالت صحت و دقت ردیابی، عملکرد ضعیف‌تری داشته است.

۴-۱۰. نتیجه‌گیری

در این فصل به بررسی نتایج حاصل از شبیه‌سازی الگوریتم ردیابی پرداخته شد. نتایج نشان می‌دهد الگوریتم ما با دقت ۷۱/۵ و صحت ۳۶/۷ عملکرد خوبی به نسبت به سایر الگوریتم‌های پیشین برجا گذاشته است. همچنین فشرده‌سازی در شرایط مختلف بررسی شد و نتایج نشان می‌دهد فشرده‌سازی در صورتی که در هر دو حالت Lost و Tracking اعمال شود، الگوریتم کارایی لازم را دارا می‌باشد، همچنین در صورتی که فشرده‌سازی در حالت Lost انجام نشود حتی کارایی آن به نسبت به زمانی که در هیچ یک از حالات فشرده‌سازی انجام نشود، پایین‌تر می‌آید. در بخش پایانی از میزان فشرده‌سازی صحبت شد که نتایج نشان می‌دهد بهترین مقدار فشرده‌سازی ۵۰ می‌باشد یعنی نهایتاً تعداد داده‌های تنک در قسمت ویژگی‌ها، ابعاد ۵۰ به خود داشته باشد تا بهترین نتیجه حاصل شود در غیر اینصورت ممکن است به علت فشرده‌سازی بیشتر یا کمتر، الگوریتم مدل‌های صحیحی برای ردیابی از خود ارائه ندهد.

فصل پنجم

نتیجه گیری

۵-۱. جمع بندی و نتیجه گیری

چالش های بسیاری در مسائل ردیابی یک شیء، وجود دارد ولی عمده ترین چالش ها عبارتند از:

۱-دقت

۲-حافظه

۳-سرعت

دقت:

دقت یکی از مهم ترین موضوعات در ردیابی یک شیء (جسم یا انسان) می باشد. اگر به درستی استخراج ویژگی صورت نگیرد، مدل درستی از هدف در نظر گرفته نمی شود در نتیجه شناسایی و ردیابی جسم را، دچار مشکل می کند و ممکن است این مدل، در برابر تغییرات ناگهانی، مانند تاری تصویر و نویز مقاوم نباشد و رانش صورت گیرد.

حافظه:

داده های ورودی به عنوان پایگاه داده، در محلی به نام حافظه سیستم ذخیره می شوند. اطلاعات دائماً در حال به روزرسانی هستند. اگر همه اطلاعات را بخواهیم در حافظه ذخیره کنیم با مشکل فضا روبرو خواهیم شد. از طرفی بخش عمده ای از داده های ورودی، به عنوان داده های زائد هستند و استفاده نخواهند شد. برای جلوگیری از سربارگیری حافظه، می توان تنها آن بخش از داده های که لازم است را در حافظه ذخیره کرد.

سرعت:

یکی از چالش‌های عمده در مسائل مربوط به ردیابی، برقرار کردن تعادل بین دقت و سرعت و حافظه است. همواره از الگوریتم‌ها و روش‌هایی در ردیابی استفاده می‌شود که علی‌رغم دقت بالا، سریع‌تر از بقیه بتوانند هدف را شناسایی و ردیابی کنند.

در ردیابی زمان واقعی، چون داده‌ها دائما به روزرسانی می‌شود در نتیجه به فضای حافظه بسیار بالایی نیاز است و چنین فضایی در ردیابی‌های چند شیء غیرممکن است. همچنین باعث کندی در سرعت ردیابی می‌شود.

برای اینکه مشکل حافظه و سرعت در ردیابی زمان واقعی تا حدود زیادی حل شود، تنها چند فریم اخیر را برای ردیابی انتخاب می‌کنیم و باقی فریم‌های قبلی را در نظر نمی‌گیریم. انتخاب اینکه تا چند فریم قبلی را در نظر بگیریم بسیار مهم است چون ممکن است ردیابی باعث بروز رانش شود.

در بالا گفته شده که، در ردیابی، همواره با سه چالش اصلی دقت، حافظه و سرعت روبرو هستیم. همچنین ذکر شده که ردیابی زمان واقعی بهتر از ردیابی آفلاین است چون در ردیابی زمان واقعی اطلاعات دائما به روزرسانی می‌شود و دقت بالاتری در تشخیص و ردیابی دارا می‌باشد. اما ردیابی زمان واقعی با مشکل حافظه و سرعت روبرو است برای حل این مسائل، تنها چند فریم اخیر را در نظر می‌گیریم. اما با همه‌ی این تفصیلات، مشکل حافظه و سرعت بطور کامل بهبود نیافته است همچنین به علت انتخاب محدودی از فریم‌های اخیر، ممکن است ردیابی دچار رانش شود. کاهش افزودنی یعنی کوچک شدن اندازه فریم‌ها به طوری که اطلاعات از بین نرود. می‌توان با استفاده از فشردگی افزودنی بخشی از اطلاعات پس-زمینه را حذف کرد در نتیجه مشکل حافظه و سرعت بسیار بهبود می‌یابد. اما همچنان مشکل رانش باقی می‌ماند. در استخراج ویژگی‌ها، می‌توان ویژگی‌های را انتخاب کرد که بیشتر در مقابل نویز مقاوم باشد، این کار توسط فشردگی داده‌های تنک انجام می‌شود که هم باعث می‌شود مشکل رانش به حداقل برسد هم باعث بهبود سرعت و حافظه شود.

یادگیری تقویتی راهی برای آموزش عامل‌ها برای انجام یک عمل از طریق دادن پاداش و تنبیه است بدون اینکه لازم باشد نحوه انجام عمل را برای عامل مشخص نمائیم.

در واقع یادگیری تقویتی یک یادگیری بدون نظارت می‌باشد که این یادگیری براساس سیاست پاداش و تنبیه انجام می‌گیرد. در ردیابی‌های زمان واقعی یکی از مشکلات اصلی رانش است. در یادگیری تقویتی با اعمال سیاست‌های مختلف، ویژگی‌هایی انتخاب می‌شود که حداکثر صحت ردیابی را تضمین کند. در نتیجه می‌توان دقت را به حداکثر رساند. برای هوشمند کردن انتخاب ویژگی، از یادگیری تقویتی استفاده می‌کنیم. اما یادگیری تقویتی علی‌رغم اینکه هوشمند است، بهینه نیست و ممکن است بهترین مسیر برای ردیابی انتخاب نکند. براین اساس از داده‌های تنک و فشرده‌سازی استفاده شده است تا بتوان، مسیرهای بهینه‌تری برای ردیابی اشیاء انتخاب شود.

این پایان‌نامه تلفیقی از یادگیری تقویتی و داده‌های تنک با دامنه فشرده است. و نتایج حاصل از آزمایشات بر روی پایگاه‌داده‌های مختلف و مقایسه با سایر الگوریتم‌های پیشین، نشان می‌دهد این الگوریتم کارایی بالاتری نسبت به همه‌ی الگوریتم‌های پیشین، در تشخیص و ردیابی اشیاء دارد. و در مقابل چالش‌های مثل نویز و انسداد اشیاء باهم، عملکرد بهتری از خود برجا گذاشته و قابلیت بالاتری در بازیابی فریم‌های از دست رفته دارد.

۵-۲. کارهای آینده

در این پایان‌نامه از الگوریتم فشرده‌سازی یکسان در هر دو حالت Trackng و Lost استفاده شده است می‌توان از الگوریتم‌های غیرهمسان استفاده کرد. همچنین می‌توان در حالت Lost از یک الگوریتم قوی تر نسبت به حالت Tracking استفاده کرد. همچنین، داده‌های ورودی با پیش فرض، غیرفشرده در نظر گرفته شدند، می‌توان الگوریتم را براساس داده‌های ورودی فشرده طراحی کرد.

- [1] حسن پور، ح.، اسدی امیری، سکینه.، مفاهیم جامع پردازش تصویر دیجیتال، فروردین ۱۳۹۴.
- [2] <http://www.sadaf-co.com>
- [3] R. M.Nasiri,Z. Wang: Perceptual aliasing factors and the impact of frame rate on video quality. In: IEEE InternationalConference on Image Processing (ICIP), pp.3476-3479 (2017).
- [4] T. Kokul, A. Ramanan, U.A.J. Pinidiyaarachchi: Online Multi-Person Tracking-by Detection Method Using ACF and Particle Filter. In: International Conference on Intelligent Computing and Information Systems (ICICIS),pp.529-536(2015).
- [5] K. Zhang, L. Zhang, and M. Yang: Real-Time Compressive Tracking. In: European Conference on Computer Vision (ECCV) 2012,pp. 866–879, (2012).
- [6] A. Kadir Jumaat, W. Eny Zarina, W.A.Rahman,A.Ibrahim, R. Mahmud: Comparison of Balloon Snake and GVF Snake in Segmenting Masses from Breast Ultrasound Images. In: Second International Conference on Computer Research and Development,pp.505 509(2010).
- [7] N. E. Tabbakha, W. H. Tan: Indoor location and motion tracking system for elderly assisted living home. In: International Conference on RoboticsAutomation and Sciences (ICORAS),pp.1-4,(2017).
- [8] K. Kim, M. Grundmann, A. Shamir, I. Matthews, J. Hodgins, I. Essa: Motion fields to predict play evolution in dynamic sport scenes. In:Computer Vision and Pattern Recognition(CVPR),pp.840-847(2010).
- [9] Wu, Yi, J. Lim, and M.H. Yang.: "Online object tracking" A Benchmark. In: IEEE Conference on Computer vision and pattern recognition (CVPR), pp.2411-2418(2013).

[10]G.D.Ramaraj, S. Venkatakrisnan, G.Balasubramanian, S. Sridhar: Aerial surveillance of public areas with autonomous track and follow using image processing.In: International Conference on Computer and Drone Applications (IConDA),pp.92-95 (2017).

[11] Litong Fan, Zhongli Wang, et al. "A survey of appearance models in visual object tracking." ACM transactions on Intelligent Systems and Technology (TIST) ,Vol.58,pp.1855-1862(2013).

[12] Hu, Weiming, et al. "A survey on visual surveillance of object motion and behaviors." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), pp.334-352 (2004).

[13]R. S. Sutton and A. G. Barto: Reinforcement Learning:An Introduction, Second edition, in progress(2017).

[14] فنبری سرخی، ع.، حسن پور، ح.، ردیابی و شناسایی مجدد افراد در شبکه‌ای از دوربین‌ها با دید مجزا بر مبنای سیستم فازی در محیط‌های بسته، مجله انجمن مهندسين برق و الکترونیک ایران-سال سیزدهم-شماره دوم-تابستان ۱۳۹۵.

[15] Mei, X., Ling, H.: Robust visual tracking and vehicle classification via sparse representation. Pattern Analysis and Machine Intelligence (PAMI) 33, 2259–2272 (2011).

[16] Ross, D., Lim, J., Lin, R., Yang, M.-H.: Incremental learning for robust visual tracking. International Journal of Computer Vision (IJCV)77, pp.125–141 (2008).

[17] Y. Xiang, A.Alahi,S.Savarese: Learning to Track: Online Multi-Object Tracking by Decision Making. In: International Conference on Intelligent Computing and Information Systems (ICICIS),pp.801-809,(2015).

[18] H. Li, C. Shen, and Q. Shi: Real-time Visual Tracking Using Compressive Sensing. In: European Conference on Computer Vision (ECCV) 2012, Part III, Live National Cricket Score(LNCS)7574, pp. 866–879,(2012).

[19] علائی، م.، امینداور، ح. ر.، " روشی جدید در بازشناسایی خودکار اهداف متحرک زمینی با استفاده از رادارهای مراقبت زمینی پالس-دابلر، " مجله انجمن مهندسين برق و الكترونیک ایران- سال ششم- شماره دوم- پاییز و زمستان ۱۳۸۸.

[20] S.C. Zhu, T.S.Lee and A.L.Yuille: Region competition: unifying snakes, region growing, energy/Bayes/MDL for multi-band image segmentation. In: Computer Vision, 1995. Proceedings., Fifth International Conference on, 884-900 (1996).

[21] Parragios.N, Deriche.R: Geodesic active regions and level set methods for supervised texture segmentation. In: Computer Vision The Proceedings of the Seventh IEEE International Conference on, pp.223-247 (1999).

[22] زرمهری، ص.، نصری، س.، معرفی و پیاده سازی الگوریتم شار نوری جهت ردیابی اشیاء، مجله انجمن مهندسين برق و الكترونیک ایران- سال یازدهم- شماره دوم- تابستان ۱۳۹۳.

[23] Kalman,R.E., "A new approach to linear filtering and prediction problems," In: American Society of Mechanical Engineers (ASME), pp. 167 – 179 (2001).

[24] Doucet,A., Godsill,S., Andrieu,C : on sequential monte carlo sampling method for bayesian filtering. In: statistics and computing, pp.197-208 (2000).

[25] M.D.Rincon, O.Urunucla, H.Jaraba "An efficient particle filter for color-based tracking in complex scenes" In Advanced Video and Signal Based Surveillance, pp.176-181 (2007).

[26] Hu,M., "Visual pattern recognition by moment invariants." IRE Transactions on Information Theory, vol.8, pp.179-187 (1962).

[27] جعفری، م.ن.، رضایی، م.، حدادنیا، ج.، شناسایی وردیابی انسان از روی تصاویر استریو، اولین همایش تخصصی علوم، فناوری و سامانه‌های مهندسی برق-مهرماه ۱۳۹۲.

[28] کرمیانی، ع.، فرج زاده، ن.، خانی، ح.، ردیابی دقیق اشیاء متحرک بر اساس اطلاعات حرکت و الگوریتم K-means اتوماتیک، بیستمین کنفرانس ملی سالانه انجمن کامپیوتر ایران- اسفند ۱۳۹۳.

- [29] X. Wang, G. Hua, and T. X. Han: Discriminative Multiple Target Tracking, pp. 306-311, (2010).
- [30] Jepson, A., Fleet, D., Maraghi, T.: Robust online appearance models for visual tracking. *Pattern Analysis and Machine Intelligence (PAMI)* Vol. 25, pp. 1296–1311 (2003).
- [31] Hare, S., Saffari, A., Torr, P.: Struck: structured output tracking with kernels. In: *International Conference on Computer Vision (ICCV)*, pp. 263–270 (2011).
- [32] Kalal, Z., Matas, J., Mikolajczyk, K.: P-n learning: bootstrapping binary classifier by structural constraints. In: *Computer Vision and Pattern Recognition (CVPR)*, pp. 49–56 (2010).
- [33] Y. Bar-Shalom and T. Fortmann, “Tracking and Data Association,” Academic Press, pp. 705-803 (1988).
- [34] Y. Boykov and D. Huttenlocher, “Adaptive Bayesian Recognition in Tracking Rigid Objects,” in: *Proc. IEEE Conference Computer Vision and Pattern Recognition*, pp. 697-704, (2000).
- [35] H. Shuo, W. Na and S. Huajun, "Object tracking method based on surf," *AASRI Procedia*, vol. 3, pp. 351-356, (2012).
- [36] A. Yilmaz, O. Javed and M. Shah, "Object tracking: A survey," *Acme computing surveys (CSUR)*, vol. 38, no. 4, pp. 1-45, (2006).
- [37] R. Rosales and S. Sclaroff, “3D Trajectory Recovery for Tracking Multiple Objects and Trajectory Guided Recognition of Actions,” in: *Proc. IEEE Conference Computer Vision and Pattern Recognition*, pp. 117-123, (1999)
- [38] G. Kitagawa, “Non-Gaussian State-Space Modeling of Nonstationary Time Series,” *Journal of the American Statistical Association*, Vol. 82, pp. 1032-1063, (1987).
- [39] Y. Bar-Shalom, *Tracking and data association: Academic Press Professional*, pp. 1202-1207, (1987).

[40] J Fan, A. K. Elmagarmid: Statistical approaches to tracking-based moving object extraction. In: Information Intelligence and Systems, pp.375-381(1999).

[41] مرآتی, امین و شهره کسایی, ردیابی اشیا متحرک غیرصلب با جابجایی میانگین مبتنی بر هسته و حذف پس زمینه, دهمین کنفرانس سالانه انجمن کامپیوتر ایران, تهران, انجمن کامپیوتر ایران, مرکز تحقیقات مخابرات ایران, 1383.

[42] B. Li and R. Chellappa, "Simultaneous Tracking and Verification via Sequential Posterior Estimation," in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, Vol. 2, pp. 110-117, (2000).

[43] P. Fieguth and D. Terzopoulos, "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, (1997).

[44] D. Comaniciu, V. Ramesh and P. Meer, "Kernel-based object tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.25, pp.564–575,(2003).

[45] D. Comaniciu ,P. Meer, "Mean shift analysis and applications," in: Proc. IEEE International Conference on Computer Vision (ICCV), Vol. 2, pp.1197–1203,(1999).

[46] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp. 603–619, (2002).

[47] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," ACM Computing Surveys, Vol. 38, No. 4, pp. 1–45, (December 2006).

[48] S. Avidan, "Support Vector Tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, pp.103-108, (August 2004).

[49] M.Pokheriya, D. Pradhan, Object detection and tracking based on silhouette based trained shape model with Kalman filter.In: Recent Advances and Innovations in Engineering (ICRAIE),pp.1-4,(2014).

[50] رحمنی، مرتضی، فضلی، سعید، محمدی، شهرام، تشخیص و ردیابی و طبقه‌بندی وسایل نقلیه با تکنیک پردازش تصویر برای سیستم با حمل و نقل هوشمند، یازدهمین کنفرانس سالانه انجمن کامپیوتر ایران، زنجان، انجمن کامپیوتر ایران، مرکز تحقیقات مخابرات ایران، ۱۳۹۱.

[51] آهنگری، فرناد، عطاران، هوبخت، برازنده، امین، ردیابی انسان توسط تفریق پس‌زمینه‌ها به روش فازی، همایش ملی علوم و مهندسی کامپیوتر، مشهد، ۱۳ آذر ۱۳۹۳.

[52] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, pp.886–893,(2005).

[53] A. Albiol, D. Martin, J. Sastre, "Face recognition using HOG-EBGM pattern Recognition Letters", Vol.29, pp.1537-1543,(2010).

[54] O. Déniz, G. Bueno, "Face recognition using Histograms of Oriented Gradients" Pattern Recognition Letters, vol. 32, pp. 1598–1603, (2011).

[55] N. Dalal and B. Triggs, "Histograms of oriented gradients for, human detection," Proc. of the IEEE Conf. on Computer Vision, and Pattern Recognition, vol. 2, pp. 886–893,(2005).

[56] <http://www.matlab1.ir>/الگوریتم-adaboost,(2017).

[57] H. Pirsivash, D. Ramanan, and C. C. Fowlkes. Globally-optimal greedy algorithms for tracking a variable number of objects. In CVPR, pages 1201–1208, (2011).

[58] S.-H. Bae and K.-J. Yoon. Robust online multi-object tracking based on tracklet confidence and online discriminative appearance learning. In CVPR, pp. 1218–1225, (2014).

[59] C. Dicle, O. I. Camps. The way they move: Tracking multiple targets with similar appearance. In ICCV, pp. 2304-2311,(2013).

[60] J H. Yoon, M.-H. Yang, J. Lim, and K.-J. Yoon: Bayesian multiobject tracking using motion context from multiple objects learning. In WACV, pp. 33–40,(2015).

- [61] A. Milan, S. Roth, and K. Schindler. Continuous energy minimization for multitarget tracking. Vol. 36 , pp. 58–72,(2014).
- [62] Mao, Lin, Yehao Xu, Fan Cheng, and Rubo Zhang. "Multi-part pedestrian tracking algorithm with fuzzy decision for partial occlusion." In 2018 Chinese Control And Decision Conference (CCDC), pp. 937-940. IEEE, (2018).
- [63] Fu, Limei. "Particle Filter Pedestrian Tracking Algorithm Based on Selected Region RGB Histogram." In 2018 International Conference on Robots & Intelligent System (ICRIS), pp. 287-290. IEEE, (2018).
- [64] Lei, Yi, and Zhijie Huang. "Research on Pedestrian Detection Algorithm Based on Monocular Vision." In 2018 International Conference on Robots & Intelligent System (ICRIS), pp. 161-163. IEEE, (2018).

Abstract

In this thesis, the identification and tracking of pedestrians in video cameras are intensively and in real time. Some pedestrian tracking applications can be used to identify a particular person or to track down criminals across the city. One of the major challenges of real-time tracking is the need for a lot of memory because the data should be updated continuously. One solution to this is to select the current frame and a limited number of subsequent frames. The problem with this method is the lack of sufficient data for training.

In such cases, a poor educational model is formed with low data, which, in conditions like brightness changes, variable displacement of the target angle, image blur and blockage are not robust and cause tracing of objects. As long as the education model is dependent on the data and, accordingly, the models are trained, there is a possibility of drift and the wrong tracking. In this thesis, the non-adaptive model and the space structure of features are used for training. Reinforcement learning is a kind of uncontrolled training combined with testing and error and reward methods. Reinforcement learning always uses multiple paths to identify and track targets, and is obstructive against intelligence, and can recover lost frames and re-target tracing. In this thesis two stages of compression on raw data and compression of thinner data have been used. Initial compression reduces computational complexity and, as a result, saves memory usage. For thrust to reach the minimum, compression on thinning data is used, which always selects the most suitable data for tracking, and does not consider the least important features. And makes it possible to provide a strong and accurate model for training. The results show that the proposed algorithm has high efficiency against noise and severe obstruction, so that the accuracy of tracking on the database was 71.5% and 36.7%, respectively.

Keywords: object tracking, Reinforcement learning, compression, thinness data, adaptive and non-conformal model.



Faculty of Computer Engineering
M.Sc. Thesis in Artificial Intelligence Engineering

Real-Time Compressive Pedestrian Tracking

Author:

By:Sajjad Jafari

Supervisors:

Dr.H.Hassanpour

Dr.V.Abolghasemi

Sep 2018