





دانشکده مهندسی کامپیوتر و فن آوری اطلاعات
گروه هوش مصنوعی

دنبال کردن چندین هدف بصورت همزمان در یک رشته ویدیویی با استفاده از تحلیل سلسله مراتبی ابرگراف غیر جهت دار

نگارش:

حمید علیمحمدی

استاد راهنما:

دکتر علی اکبر پویان

اساتید مشاور:

دکتر مرتضی زاهدی

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

شهریور ۱۳۹۴

دانشگاه شاهرود

دانشکده: مهندسی کامپیوتر و فن آوری اطلاعات

گروه: هوش مصنوعی

پایان نامه کارشناسی ارشد آقای حمید علیمحمدی

تحت عنوان: دنبال کردن چندین هدف بصورت همزمان در یک رشته ویدیویی با استفاده از گراف های سطح بالای سلسله مراتبی هدایت نشده

در تاریخ توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد مورد ارزیابی و با درجه عالی مورد پذیرش قرار گرفت.

امضاء	اساتید مشاور	امضاء	اساتید راهنما
	نام و نام خانوادگی: دکتر مرتضی زاهدی		نام و نام خانوادگی: دکتر علی اکبر پویان

امضاء	نماینده تحصیلات تکمیلی	امضاء	اساتید داور
	نام و نام خانوادگی: مهندس فرهادی		نام و نام خانوادگی:
			نام و نام خانوادگی:

تقديم به

خدایا، به من زیستنی عطا کن که در لحظه مرگ بر بی‌ثمری لحظه‌ای که برای زیستن گذشته است، حسرت نخورم و مردنی عطا کن که بر بیهودگی‌اش سوگوار نباشم. عمر ما آنقدر طولانی نیست که مسیر زندگی را یک بار برای کسب تجربه بپیماییم و بار دیگر برای به کار بردن تجربه‌ها. در زندگی یا باید دل به دریا زده و با هراس قدم در جاده زندگی بگذاری-مسیری که در طول آن نه راهنمایی حضور دارد نه چراغی-و یا ابتدا خود را به چراغ روشن آگاهی مجهز نمایی و سپس با ایمان و اطمینان پا در راه بگذاری.

تعهد نامه

اینجانب حمید علیمحمدی دانشجوی دوره کارشناسی ارشد رشته مهندسی کامپیوتر-گرایش هوش مصنوعی دانشکده مهندسی کامپیوتر و فن آوری اطلاعات دانشگاه شاهرود نویسنده پایان نامه " دنبال کردن چندین هدف بصورت همزمان در یک رشته ویدیویی با استفاده از گرافهای سطح بالای سلسله مراتبی هدایت نشده " تحت راهنمایی دکتر علی اکبر پویان متعهد می شوم.

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

تاریخ

امضای دانشجو

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) متعلق به دانشگاه شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

امروزه روش‌های مختلف رهگیری در یک فایل ویدیویی در قالب دو دسته کلی مطرح می‌شوند. دسته اول مربوط به مسائلی هستند که فرآیند رهگیری را بصورت دنبال کردن یک هدف در یک رشته ویدیویی تعریف می‌کنند و دسته دوم روش‌هایی را مورد بررسی قرار می‌دهد که توانایی رهگیری چندین هدف را بصورت همزمان دارا می‌باشند. هدف رهگیری همزمان چند شی، تخمین وضعیت آنها با حفظ هویت و با در نظر گرفتن تغییرات حرکت و ظاهر آنها با گذشت زمان می‌باشد. در یک محیط مختلط و شلوغ، این موضوع به خصوص بواسطه تصادم پی در پی ناشی از بهم ریختگی، ظاهر شبیه به یکدیگر سایر اهداف و بیشتر چالش برانگیز می‌شود.

دنبال کردن عابران پیاده در صحنه‌های واقعی یک مولفه بحرانی برای بسیاری از کاربردهاست، از قبیل سیستم‌های نظارت ویدیویی و سیستم‌های دستیار راننده. شناسایی و طبقه‌بندی دقیق عابران در بهبود فرآیند دنبال کردن عابران موثر است. اخیراً الگوریتم‌های دنبال کردن عابران پیاده توانسته اند کارایی خوبی را در صحنه های کم ترد بدست آورند اما کارایی آنها در صحنه های شلوغ بشدت کاهش می یابد بطوریکه ابهام در شکل ظاهری به سبب تصادم‌های زیاد ارتباط فضایی پیچیده ای را در بین عابران مختلف ایجاد می‌کند.

از طرف دیگر الگوریتم‌های دنبال کننده در فضای حالت مانند کالمن و فیلتر ذره‌ای توانایی مقابله با تصادم‌های طولانی مدت و حل مسئله‌ای با پیچیدگی بالا و شباهت ظاهری زیاد اهداف را ندارند.

در این تحقیق مسئله دنبال کردن عابران پیاده در یک محیط شلوغ بوسیله یک ابر گراف مدل شده است. یک استراتژی جستجوی محلی به کل برای طبقه بندی شناسایی‌ها بصورت سلسله مراتبی بکار گرفته شده است که سبب کاهش بسیار زیاد پیچیدگی‌های محاسباتی در هنگام مواجهه با

تغییرات ناگهانی ظاهری و حرکتی اهداف می‌شود. نخست فایل ویدیوی به چند قسمت در حوزه زمان تقسیم می‌شود و همسایگی‌های متراکم در هر قسمت جستجو می‌شوند تا چندین مسیرک^۱ طولانی تر ساخته شوند. سپس قسمت‌های نزدیک بهم برای ساختن تقسیم بندی جدید جهت بکارگیری در لایه بعد باهم یکپارچه می‌شوند. این دو مرحله تکرار می‌شوند تا زمانی که فقط یک قسمت در لایه باقی بماند یعنی کل ویدیو مورد نظر و جستجوی همسایگی‌های متراکم در این قسمت انجام می‌شود تا مسیر نهایی حرکت اهداف بدست آید. این روش قادر است تا بطور همزمان چند هدف را در یک محیط شلوغ دنبال کند. نتایج بدست آمده بر روی پایگاه داده ParkingLot بخوبی نشان می‌دهد که مسیر حرکت اهداف بصورت صحیح دنبال شده است.

کلمات کلیدی:

رهگیری، آشکارسازی، تجمیع کلی، سلسله مراتبی، مسیرک

^۱ Tracklet

فهرست مطالب

۱- مقدمه و کلیات	۱
۱-۱- پیشگفتار	۲
۲-۱- طرح مساله	۲
۳-۱- هدف و نحوه رویکرد پژوهش	۴
۴-۱- ساختار پایان نامه	۵
۲- ادبیات موضوع و بررسی پژوهشهای مرتبط	۶
۱-۲- مقدمه	۷
۲-۲- ادبیات موضوعی	۷
۳-۲- مراحل اصلی رهگیری هدف	۱۳
۳-۲-۱- آشکارسازی هدف:	۱۴
۳-۲-۱-۱- آشکار سازی نقاط مهم	۱۵
۳-۲-۱-۲- تفریق پس زمینه	۱۵
۳-۲-۱-۳- بخش بندی تصویر	۱۶
۳-۲-۱-۴- یادگیری با سرپرست	۱۶
۳-۲-۲- طبقه بندی اهداف	۱۷
۳-۲-۲-۱- طبقه بندی بر مبنای شکل:	۱۷
۳-۲-۲-۲- طبقه بندی بر اساس حرکت:	۱۸
۳-۲-۲-۳- طبقه بندی مبتنی بر رنگ:	۱۸
۳-۲-۲-۴- طبقه بندی بر مبنای بافت:	۱۹
۳-۲-۳- رهگیری هدف	۱۹
۳-۲-۳-۱- روشهای رهگیری مبتنی بر نقطه	۱۹
۳-۲-۳-۱-۱- کالمن فیلتر	۲۰
۳-۲-۳-۱-۲- فیلتر ذرات	۲۰
۳-۲-۳-۲- رهگیری چند فرضی	۲۱
۳-۲-۳-۲- روشهای رهگیری مبتنی بر کرنل	۲۱

۲۱.....	تطبیق الگوی نمونه.....	۱-۲-۳-۳-۲
۲۲.....	روش جابجایی میانگین.....	۲-۲-۳-۳-۲
۲۲.....	ماشین بردار پشتیبان.....	۳-۲-۳-۳-۲
۲۳.....	رهگیری مبتنی بر لایه.....	۴-۲-۳-۳-۲
۲۳.....	رهگیری سایه نما.....	۲-۳-۳-۳-۳
۲۳.....	رهگیری با استفاده از کانتور.....	۱-۳-۳-۳-۲
۲۴.....	تطبیق شکل.....	۲-۳-۳-۳-۲
۳- تحلیل تئوریک الگوریتمهای فیلتر کالمن و فیلتر ذره‌ای.....		
۲۶.....	مقدمه.....	۱-۳
۲۶.....	فیلتر کالمن.....	۲-۳
۳۱.....	فیلتر ذره‌ای.....	۳-۳
۳۵.....	علل نیاز به رویکرد نوین.....	۴-۳
۴- مدل کردن مسئله رهگیری چند هدف با ابر گراف.....		
۳۷.....	مقدمه.....	۱-۴
۳۹.....	مروری بر تئوری ابر گراف.....	۲-۴
۴۰.....	برقراری ارتباط میان دو نوع متفاوت از موجودیتها.....	۱-۲-۴
۴۵.....	ارتباط سلسله مراتبی ابر گراف بر مبنای دنبال کننده.....	۴-۳
۴۷.....	ساختار گراف وابستگی غیر جهت دار.....	۴-۴
۴۸.....	شباهت (وابستگی) ظاهری.....	۱-۴-۴
۴۹.....	شباهت حرکتی.....	۲-۴-۴
۵۰.....	شباهت همواری مسیر.....	۳-۴-۴
۵۱.....	جستجوی همسایگی های متراکم مسیر کها.....	۴-۵
۵۴.....	روش پیشنهادی جهت بهبود الگوریتم.....	۶-۴
۵۴.....	کنترل تصادم با اضافه کردن راس فرضی.....	۱-۶-۴
۵۴.....	نتایج پیاده سازی.....	۷-۴

۵- نتیجه گیری و پیشنهادات..... ۵۸

۵-۱- جمع بندی و نتیجه گیری..... ۵۹

۵-۲- پیشنهادات برای ادامه تحقیق..... ۶۲

مراجع..... ۶۳

فهرست شکل ها

- شکل ۴-۱: نحوه عملکرد الگوریتم پیشنهادی بر اساس تحلیل سلسله مراتبی ابر گراف..... ۳۸
- شکل ۴-۲: ساختار کلی ابر گراف..... ۴۱
- شکل ۴-۳: ابر گرافهای جهتدار..... ۴۲
- شکل ۴-۴: مدل کردن مدار با هایپر گراف..... ۴۲
- شکل ۴-۵: توپولوژی توصیفی برای یک مسئله RDF..... ۴۴
- شکل ۴-۶: مدل کردن مسئله RDF با یک ابر گراف..... ۴۴
- شکل ۴-۷: بازیابی مسیرکها در یک فرآیند بهینه سازی از لایه سوم به لایه دوم..... ۴۷
- شکل ۴-۸: تصحیح خطای خط سیر اهداف در هنگام مخفی شدن عابر..... ۵۴
- شکل ۴-۹: نتایج آنالیز بخش اول پایگاه داده ParkingLot..... ۵۶

فصل اول

مقدمه و کلیات

۱-۱- پیشگفتار

بینایی ماشین یکی از زمینه‌های پراهمیت علم کامپیوتر است و سال‌هاست که محققان به پیشرفت و توسعه آن پرداخته‌اند. مهمترین هدف بینایی ماشین شبیه‌ساز عملکرد سیستم بینایی انسان است و شاید به جرات بتوان گفت که بزرگترین هدف بینایی ماشین ساختن یک ماشین دارای بینایی همه منظوره است. از چنین ماشینی انتظار می‌رود که بتواند خود را در محیط‌های دلخواه هدایت کند و اشیاء را در مسیر خود شناسایی نماید و به مقصدی که باید برسد، خود را برساند. در واقع امروزه بینایی ماشین از بینایی انسان فاصله زیادی دارد و هنوز جای تحقیق و مطالعه بسیاری باقی است. سیستم بینایی انسان با دارا بودن توانای آنالیز و تحلیل تصاویر به روابط فیزیکی موجود در صحنه پی می‌برد. هدف نهایی همه الگوریتم‌ها و روش‌های پیشنهاد شده در تحقیقات، ایجاد یک ماشین دارای بینایی و هوش لازم برای تفسیر اطلاعات بینایی است. چنین هوشی در واقع قابلیت پردازش تصاویر اشیاء یا نواحی خاصی از تصاویر را دارد. در این میان اغلب اشیاء موجود در تصویر دارای حرکت می‌باشند. از این روی یکی از مسائل مهم در بینایی ماشین بررسی حرکت اشیاء موجود در صحنه می‌باشد که می‌بایست حرکت آنها را دنبال و تحلیل نمود.

دنبال کردن یک هدف در تصویر را می‌توان بصورت پیدا کردن موقعیت شیء مورد نظر در تک‌تک فریم‌های یک رشته ویدیویی در نظر گرفت. سیستم دنبال کننده شیء بسته به کاربرد آنها ممکن است اطلاعات دیگری مانند شکل شیء، جدا شدن بخش یا چیزی از آن و... را نیز مشخص نماید.

۱-۲- طرح مساله

امروزه روش‌های مختلف رهگیری در قالب دو دسته کلی مطرح می‌شوند. دسته اول مربوط به مسائلی هستند که فرآیند رهگیری را بصورت دنبال کردن شیء مورد نظر در یک رشته ویدیویی تعریف می‌کنند و دسته دوم روش‌هایی را مورد بررسی قرار می‌دهد که توانایی رهگیری چندین هدف

را بصورت همزمان دارا می‌باشند. هدف رهگیری همزمان چند شی در یک ویدیو، تخمین وضعیت آنها با حفظ هویت و در نظر گرفتن تغییرات حرکت و ظاهر آنها با گذشت زمان می‌باشد. در یک محیط مختلط و شلوغ، این موضوع به خصوص بواسطه تصادم پی در پی ناشی از بهم ریختگی، ظاهر شبیه به یکدیگر سایر اهداف و بیشتر چالش برانگیز می‌شود. از جمله مهم‌ترین روش‌ها در دسته اول می‌توان به روش اول انتقال میانگین، فیلتر کالمن و فیلتر ذره‌ای اشاره کرد. نکته مهم در این زمینه این است که روش‌هایی مانند فیلتر کالمن و فیلتر ذره‌ای با تغییراتی در فرآیند رهگیری آنها به الگوریتم‌هایی با توانایی دنبال کردن چند هدف بصورت همزمان تبدیل می‌شوند.

کالمن فیلتر روشی است که از مجموعه‌ای از اندازه‌گیری‌های مشاهده شده در طول زمان استفاده می‌کند. این اندازه‌گیری‌ها می‌تواند شامل نویز و یا سایر خطاها باشد. کالمن فیلتر تخمینی از متغیرها تولید می‌کند که می‌تواند بسیار دقیق‌تر از تخمینی باشد که بر مبنای یک اندازه‌گیری به تنهایی بدست آید. به عبارت دیگر، کالمن فیلتر به طور بازگشتی بر روی جریانی از ورودی‌های آغشته به نویز عمل می‌کند و می‌تواند تخمین آماری بهینه‌ای از وضعیت سیستم مورد مطالعه تولید کند.

دنبال کننده دیگری که در بسیاری از مطالعات مورد توجه قرار گرفته است، دنبال کننده مبتنی بر فیلتر ذره‌ای می‌باشد. این فیلتر از راهکار فضای حالت برای مدل کردن سیستم استفاده می‌کند. در این فیلترها تعریف مسئله دنبال کردن شی بصورت تعیین حالت و همچنین چگونگی پیش‌بینی موقعیت هدف و بروزرسانی به روش بازگشتی انجام می‌پذیرد. فیلتر ذره‌ای تکنیکی بر پایه تخمین بیز بازگشتی برای حالت مجهول سیستم می‌باشد. از جمله دلایل محبوبیت فیلتر ذره‌ای می‌توان به توانایی دنبال کردن چند فرض، توانایی کار کردن با مدل‌های غیر خطی و امکان پیاده‌سازی آن بصورت بلادرنگ و موثر اشاره کرد.

با این وجود اکثر روش‌های قبلی از جمله فیلتر ذره‌ای دنبال کردن شی را با توجه به میزان شباهت ظاهری و الگوی حرکت انجام می‌دهند. اینگونه روش‌ها در حین رهگیری یک هدف بسته در

یک صحنه شلوغ بدلیل عدم کنترل تصادم و مسئله مخفی شدن هدف با مشکلات زیادی روبرو می-شوند.

روش مورد نظر در این پایان نامه بکارگیری سلسله مراتبی ابرگراف‌های غیر جهت دار است. این روش، دنبال کردن شی را بعنوان یک مسئله جستجوی سلسله مراتبی همسایگی‌های متراکم در نظر گرفته و بصورت پویا گراف غیرجهت‌دار هدف را ایجاد می‌نماید. مدلسازی موجود امکان رهگیری چندین هدف را بصورت همزمان در یک محیط شلوغ با تصادم زیاد فراهم می‌کند.

۱-۳- هدف و نحوه رویکرد پژوهش

در سال‌های اخیر به دلیل توسعه روش‌های آشکارسازی و تشخیص مطمئن اشیا حتی در محیط‌ها و صحنه‌های شلوغ، عملکرد روش‌های رهگیری توسط آشکارسازی بهبود قابل توجهی داشته است [۱و۲]. به طور کلی، روش‌های رهگیری توسط آشکارسازی، با استفاده از اطلاعات جمع‌آوری شده از آشکارسازها مکان هندسی^۱ مربوطه را ایجاد می‌کند. نکته قابل توجه در فرآیند رهگیری مدیریت آشکارسازی نادقیق و یا عدم وجود آشکارسازی به هنگام بروز تصادم بین اهداف بسیار مشکل می‌باشد. با در دست داشتن آشکارسازی‌های صورت گرفته برای یک مجموعه از فریم‌ها، مکان‌های هندسی کوچکتر با به هم پیوستن آشکارسازی‌ها پدید آمده و در نهایت مکان‌های هندسی کوچکتر به منظور ساخت مکان هندسی طولانی‌تر به یکدیگر پیوند داده می‌شوند. بنابراین در این دیدگاه، تجمیع کلی^۲ بسیار اثرگذار و حائز اهمیت است [۳]. به هر حال، به هنگام وقوع تصادم‌های بلند مدت و پیوسته، تشخیص و جداسازی اهداف مختلف دشوار بوده و همین موضوع در این پایان نامه بعنوان چالش اصلی مورد تجزیه و تحلیل قرار گرفته است. در واقع هدف این پایان نامه، تخمین وضعیت چندین هدف با حفظ هویت آن‌ها و با در نظر گرفتن تغییرات حرکت و ظاهر آن‌ها با گذشت زمان می‌باشد. در یک

^۱. Trajectory

^۲. Global association

محیط مختلط و شلوغ، این موضوع به خصوص بواسطه تصادم پی در پی ناشی از بهم ریختگی، ظاهر شبیه به یکدیگر سایر اهداف و بیشتر چالش برانگیز می‌شود که سعی شده است تا حدودی این چالش نیز مد نظر قرار گیرد.

۱-۴- ساختار پایان نامه

در فصل دوم ادبیات موضوع، پژوهش‌های مرتبط در این زمینه و نتایج بدست آمده توسط محققین مختلف با رجوع به متون فنی گوناگون معرفی می‌گردد. از آنجا که می‌توان روش‌های مطرح برای مساله دنبال کردن یک شی را از زوایایی مختلفی مورد بحث قرار داد، سعی شده است به جنبه‌های گوناگون مسئله در تحقیقات قبل پرداخته شود.

فصل سوم به تحلیل تئوریک دو الگوریتم مهم در فرآیند دنبال کردن پرداخته شود. در این فصل الگوریتم‌های فیلتر کالمن و فیلتر ذره‌ای مورد تجزیه و تحلیل قرار گرفته است.

در فصل چهارم روش پیاده سازی شده مورد بررسی قرار می‌گیرد و نتایج بدست آمده از آن در حالت کلی ارزیابی می‌گردد.

آنچه که در فصل پنجم ارائه خواهد شد شامل نتیجه‌گیری و جمع‌بندی حاصل از پژوهش انجام شده است. در انتها نیز نقاط ضعف و قوت روش بیان و همچنین موضوعاتی برای تحقیق در آینده پیشنهاد می‌گردد.

فصل دوم

ادبیات موضوع و بررسی پژوهش‌های مرتبط

۲-۱- مقدمه

یکی از مسائل مهم و در حال توسعه در پردازش تصویر و بینایی ماشین مساله ردیابی اشیا است. در واقع ردیابی اشیا نمایش تغییرات موقعیت یک شی و دنبال کردن آن در یک دنباله تصاویر ویدیویی با یک هدف خاص می باشد.

پیشینه ایجاد پدیده ردیابی اشیا به مسائل نظامی برمی گردد، در واقع مسائلی چون هدف یابی مخصوصا اهداف متحرک، ردیابی مسیر موشک شلیک شده برای اطمینان از صحت عمل هدف گیری و یا ردیابی مسیر موشک شلیک شده از سوی دشمن برای جلوگیری از اصابت آن به هدف و مسائلی از این قبیل سبب پدید آمدن یک مقوله جدید به نام ردیابی اشیا در پردازش تصویر نظامی گردید.

در این فصل روشهای مختلفی را که برای دنبال کردن شیء ارائه شده اند، مرور خواهیم کرد. برای این کار یک روش پایین به بالا را در تشریح مواردی که در ایجاد یک سیستم دنبال کننده شیء باید مورد توجه قرار گیرند، به کار می بریم. در اولین گام مروری بر الگوریتمهای مختلف رهگیری اشیا متحرک انجام می دهیم. در گام بعد مراحل اصلی رهگیری یک هدف شامل آشکارسازی هدف، طبقه بندی هدف و رهگیری هدف را مورد بررسی قرار می دهیم و روشهای مختلف بکار گرفته شده در هر حوزه را بیان می کنیم.

۲-۲- ادبیات موضوعی

امروزه ردیابی اشیا در زمینه های مختلف مانند زمینه های اکتشافی در حوزه هوانوردی، فشرده سازی هوشمند ویدیو، مراقبت ویدیویی، کنترل مبتنی بر بینایی ماشین، تصویر برداری پزشکی و ... کاربردهای بسیار گسترده ای یافته اند. الگوریتمهای مورد مطالعه در این حوزه می بایست قابلیت مقابله با پدیده هایی از جمله تغییر روشنایی محیط، همپوشانی اشیا، عدم حرکت اشیا با سرعت ثابت و یا در راستای یک خط مستقیم و ... را دارا باشند و از پایایی کافی نیز در طول زمان برخوردار باشند.

ارائه چنین الگوریتم‌هایی نیازمند بررسی روش‌های حرکت اشیا و مفاهیم مرتبط با آنها، مسئله حرکت نسبی شی، توابع و روابط ریاضی هندسی حاکم بر الگوهای حرکت، تخمین‌های آماری لازم برای آشکارسازی‌ها و تشخیص نوع شی متحرک و ... می‌باشد که در ادامه نگاهی به روش‌های مختلف دنبال کردن شی می‌پردازیم. لیپتون و همکاران برای استخراج بخش‌های متحرک یک تصویر، تفاضل بین دو فریم را به عنوان راهکار ارائه کرده‌اند [۴]. از طرفی استافر و همکاران یک مدل ترکیبی گوسی بر مبنای مدل پس زمینه برای آشکارسازی هدف پیشنهاد کرده‌اند [۵]. در [۶] تفاضل پس زمینه تصویر مرجع با سایر تصویر به صورت پیکسل به پیکسل برای تشخیص نواحی متحرک ارائه شده است. کیم و همکاران در [۷] یک روش هابیرید که ترکیبی از تفاضل‌گیری بین سه فریم به همراه یک مدل تطبیقی برای تفریق^۱ (حذف) پس زمینه ارائه کرده‌اند. در [۸]، ترکیبی از حذف پس زمینه و اختلاف بین فریم‌ها ارائه شده که عملکرد روش‌های حذف پس‌زمینه و تفاضل بین فریم را بهبود داده است.

کولینس [۹] روشی برای رهگیری افراد در حال حرکت با استفاده از افزایش شاخص همبستگی در تصویر ارائه کرده است. بیشاپ در [۱۰] راهکاری مبتنی بر فیلتر کالمن ارائه شده که موقعیت و سرعت هدف را در دو جهت تخمین می‌زند.

از طرفی اسکالروف و همکاران در سال ۱۹۹۹ استفاده از فیلتر کالمن را به منظور تخمین مکان هندسی هدف در سه بعد گسترش داده‌اند [۱۱].

کنا و همکاران یک روش حذف پس زمینه تطبیقی ارائه کرده‌اند که در آن اطلاعات رنگ و نیز گرادیان به منظور سازگاری بیشتری با سایه و رنگ‌های کاذب و غیر قابل اطمینان ترکیب شده‌اند. سپس رهگیری به طور خلاصه در سه سطح انجام می‌شود: ناحیه، افراد و گروه‌ها هر ناحیه یک شکل بسته هست و نواحی می‌توانند با یکدیگر ادغام شده و یا بشکنند. یک فرد متشکل از یک یا چند ناحیه است که با رعایت قیود هندسی مربوط به بدن یک انسان به یکدیگر ملحق می‌شوند. یک گروه نیز متشکل از یک یا چند فرد است [۱۲].

^۱ . Subtraction

چنگ و همکاران [۱۳] در سال ۲۰۰۶ یک ویژگی بر مبنای رنگ و یک ویژگی فضایی از شیء را به منظور تعیین و شناسایی شیء که باید رهگیری شود ارائه کرده‌اند. ویژگی فضایی از مرز بسته‌ای که هدف را احاطه کرده است استخراج می‌شود. در مقابل، ویژگی‌های رنگی از مقدار میانگین و استاندارد هر هدف استخراج می‌شود. در سال ۲۰۰۷ سیز و همکاران [۱۴]، اندازه‌گیری میزان شباهت اهداف را با استفاده از معیار فاصله باهاتاچاریا^۱ پیشنهاد کرده‌اند. هر چه فاصله باهاتاچاریا کمتر باشد میزان شباهت بیشتر است.

پارک و همکاران یک روش رهگیری هدف نیمه خودکار ارائه کرده‌اند که شامل دو گام استخراج شی درون فریم و رهگیری اشیا بین فریم‌ها می‌باشد [۱۵]. با استفاده از جداسازی نواحی همگن، مداخله انسان در استخراج شی در داخل یک فریم کاهش می‌یابد. همچنین با استفاده از تخمین حرکت یک بعدی، زمان پردازش رهگیری شی از یک فریم به فریم دیگر کاهش می‌یابد. در این مرجع یک روش سیلاب^۲ بهبود یافته برای استفاده در الگوریتم جریان آب^۳ ارائه شده است. دنگ و مانجونات (۲۰۰۱) روشی برای جداسازی نظارت نشده در عکس و ویدئو ارائه کرده‌اند [۱۶]. این الگوریتم که JSEG نامیده می‌شود، بر مبنای نواحی بافت رنگی در عکس و ویدئو عمل می‌کند. این روش شامل دو مرحله کوانتیزاسیون رنگی و جداسازی فضایی می‌باشد. در گام اول، به منظور تفکیک نواحی در یک تصویر، رنگ‌های موجود در تصویر به چند گروه اصلی کوانتیزه می‌شوند. سپس، پیکسل‌ها با برچسب گروه رنگی متناظرشان جایگزین می‌شوند. نتیجه این عمل، یک تصویر تقسیم‌بندی شده می‌باشد. با یک جداسازی مناسب، به پیکسل‌های متعلق به مرزهای نواحی رنگی مختلف، مقادیر بزرگتری نسبت به پیکسل‌های داخل ناحیه اختصاص داده می‌شود. در رابطه با ویدئو یک مرحله اضافه می‌شود و این عمل برای هر فریم تکرار می‌شود.

^۱ Bhattacharya

^۲ . Flooding method

^۳ . Water shed

یاکوف و همکاران الگوریتمی برای جداسازی خودکار اهداف متحرک در ویدیوهای MPEG-4 ارائه کرده‌اند [۱۷]. MPEG-4 بر مبنای تفکیک هر فریم به صفحات شی ویدیویی^۱ (VOP) در دنباله‌ای از تصویر تشکیل شده است. هر صفحه متناظر با یک شیء متحرک در صحنه می‌باشد. پردازش اصلی، تقسیم نواحی به پس زمینه و پیش‌زمینه^۲ بر مبنای اطلاعات بدست آمده از حرکت شی در صفحه می‌باشد. مسأله جداسازی به عنوان تشخیص اشیا متحرک بر روی پس زمینه ثابت^۳ است. حرکت دوربین توسط مدل حرکت پرسپکتیو با ۸ پارامتر جبران می‌شود. در ابتدا، تقسیم‌بندی فضا توسط الگوریتم جریان آب بدست می‌آید. سپس، گرادیان Canny به منظور تخمین گرادیان فضایی در فضای رنگ به کار برده می‌شود. پس از آن، الگوریتم جریان آب^۴ بهینه شده اعمال می‌شود. بر مبنای تقسیم‌بندی اولیه، نواحی به پیش‌زمینه و پس‌زمینه تقسیم می‌شوند. حرکت هر کدام از نواحی پیش‌زمینه توسط تطابق ناحیه در قالب‌کاری سلسله مراتبی تخمین زده می‌شود. بردارهای حرکت تخمینی در مدل میدان تصادفی مارکوف (MRF) به کار گرفته می‌شود که با استفاده از منطق "ابتدا مطمئن‌ترین"^۵ بهینه‌سازی می‌شود. این عمل منجر به طبقه‌بندی اولیه ناحیه می‌شود. MRF شامل اطلاعاتی از فریم قبلی می‌باشد. گام آخر شامل یک حافظه دینامیک به منظور اطمینان از همگرایی موقت فرآیند جداسازی می‌باشد.

ونکاتش (۲۰۰۴) دیدگاهی برای تخمین خودکار تعداد اهداف و جداسازی اهداف ویدیویی متحرک از ویدیوهای MPEG-4 با استفاده از بردارهای حرکت در [۱۸] ارائه کرده است. از آنجا که در ویدیوهای MPEG فشرده، بردارهای حرکت اسپارس هستند یعنی یک بردار حرکت در هر بلوک ماکرو وجود دارد، روشی برای افزایش اطلاعات حرکت از چند فریم قبل و بعد از فریم کنونی ارائه شده است. میانمایی توسط

^۱ . Video object planes

^۲ . Foreground

^۳ . Static

^۴ . Rainfall

^۵ . Highest Confidence First (HCF)

مدین^۱ فیلتر انجام شده است بنابراین به هر پیکسل در فریم یک بردار حرکت اختصاص می‌یابد. سپس، جداسازی انجام می‌شود. از آنجا که به منظور تخمین پارامترهای حرکت اطلاعات کافی در دسترس نیست، الگوریتم حداکثر نمودن میانگین^۲ (EM) مورد استفاده قرار می‌گیرد. الگوریتمی برای تخمین تعداد مدل-های متحرک ارائه شده است. به محض اینکه جداسازی اولیه انجام شود، صفحات اشیا ویدیویی توسط رهگیری تولید می‌شود. در نهایت، مرحله اصلاح مرزها بر اشیا موجود در ویدیو اعمال می‌شود که در آن پیکسل‌های موجود در مرزها به اشیا ویدیویی صحیح اختصاص پیدا می‌کند.

هدف مزایس و همکاران (۲۰۰۴) جداسازی اشیا موجود در یک دنباله ویدئویی می‌باشد [۱۹]. در این الگوریتم سه مرحله وجود دارد: ابتدا جداسازی فریم اول با استفاده از اطلاعات رنگ، حرکت و موقعیت صورت می‌گیرد. سپس یک رهگیری موقت انجام شده و در نهایت یک فرآیند ادغام ناحیه^۳ انجام می‌شود. در مرحله جداسازی از روش K-means با قید پیوستگی استفاده می‌شود. رهگیری توسط طبقه‌بندی کننده Bayes انجام می‌شود. یک پردازش قانونمند به منظور اختصاص مجدد پیکسل‌های تغییر یافته به نواحی موجود و نیز چگونگی معرفی نواحی جدید انجام می‌شود. ادغام نواحی بوسیله یک دیدگاه مبتنی بر مکان هندسی به جای حرکت در سطح فریم صورت می‌گیرد. یکی از مزایای این روش امکان رهگیری مؤثر اهداف جدید که در صحنه ظاهر می‌شوند و نیز اهداف که به سرعت حرکت می‌کنند می‌باشد.

ویمینگ هو و همکاران در سال ۲۰۱۲ یک الگوریتم تعلیم لگاریتمی-اقلیدسی ریمانیان^۴ افزایشی ارائه کرده-اند [۲۰]. ابتدا ماتریس کوواریانس ویژگی‌های تصویر با استفاده از معیار لگاریتمی-اقلیدسی ریمانیان به یک فضای برداری نگاشت می‌شوند. اطلاعات جانمایی فضایی محلی و سراسری با استفاده از یک مدل ظاهر اقلیدسی-لگاریتمی بدست می‌آید. تعیین حالت بایاسین مبتنی بر فیلتر ذرات برای رهگیری یک و یا چند

^۱ . Median filter

^۲ . Expectation Maximization

^۳ . Region Merging

^۴ . log-Euclidean Riemannian

هدف با وجود تصادم به کار برده می‌شود. تغییر در ظاهر هدف با به روزرسانی افزایشی^۱ مدل ظاهر اقلیدسی-لگاریتمی تعیین می‌شود.

چانگ هانگ و همکاران (۲۰۱۳) روشی پیشنهاد کردند که در آن ورودی به عنوان نتایج تشخیص هدف به صورت فریم به فریم در نظر گرفته می‌شود [۲۱]. اولین مجموعه مسیرک‌های هدف توسط استراتژی دو سطحی متداول بدست می‌آید. سپس، پاسخ‌های قابل اطمینان که از آشکارسازی بدست آمده‌اند به هم متصل می‌شوند. به هم پیوستن و تجمیع مسیرهای آشکار شده که مشکوک هستند تا زمان جمع‌آوری شواهد بیشتر به تعویق می‌افتند. مسیرهای چندگانه با استفاده از تجمیع سلسله مراتبی برای دستیابی به مسأله MAP^۲ استفاده می‌شود.

رانا و همکاران (۲۰۱۳) یک روش رهگیری مقاوم برای استخراج یک هدف متحرک از یک فریم را تحت شرایطی خاص ارائه می‌کنند [۲۲]. این روش شامل دو گام است. ابتدا سه ویژگی ضعیف به منظور رهگیری کلی هدف ترکیب می‌شوند. سپس، مرزهای رهگیری کننده به منظور استخراج هدف مورد نظر تعیین می‌شوند. روش‌هایی که به تازگی در این مرجع معرفی شده‌اند شامل هیستوگرام‌های همپوشان شدت روشنایی^۳ (OHI) و یک روش جداسازی جدید که از تفکیک آنالین پس زمینه از مرز و پالس‌های ساختار یافته مبتنی بر تکه مرزهای^۴ جدا از هم استفاده می‌کنند تشکیل شده است. تکه مرزها تکه‌های گسسته‌ای از مرزها هستند. روش پنجره متحرک (لغزان) به منظور تعیین محل تقریبی (کلی) هدف استفاده می‌شوند.

شائو و همکاران (۲۰۱۳) دو نوآوری اساسی دارد: ابتدا یک الگوریتم تصمیم‌گیری مبتنی بر سطح آستانه و با استفاده از مدل چند پس زمینه^۵ برای جداسازی اهداف موجود در ویدیو ارائه شده است [۲۳]. سپس یک

^۱ . Incremental

^۲ . Maximum A Posteriori

^۳ . Overlapped High Intensity

^۴ . edglet

^۵ . Multi background

(قاب کاری/ روش)^۱ رهگیری اهداف در ویدیو مبتنی بر فیلتر ذره‌ای با تابع شباهت ارائه شده است که متشکل از میزان شباهت هیستوگرام رنگی اندازه‌گیری شده بر مبنای اندازه‌گیری فاصله انتشار و حرکت شاخص^۲ جداسازی هدف در ویدیو می‌باشد. این قاب کاری می‌تواند تغییرات ناخواسته در روشنایی و در هم ریختگی ایجاد شده پس زمینه را مدیریت کند و می‌تواند اهداف متحرک غیر صلب^۳ را نیز رهگیری کند. الگوریتم تصمیم سطح آستانه یک مقدار آستانه بهینه برای جداسازی تعیین می‌کند. هیستوگرام مبتنی بر رنگ نیز برای رهگیری بهتر اهداف غیر صلب استفاده می‌شود. در روش ارائه شده در این مرجع یک هیستوگرام رنگی تک بعدی به جای نوع سه بعدی آن استفاده شده است که در نتیجه پیچیدگی محاسباتی می‌تواند کاهش یابد.

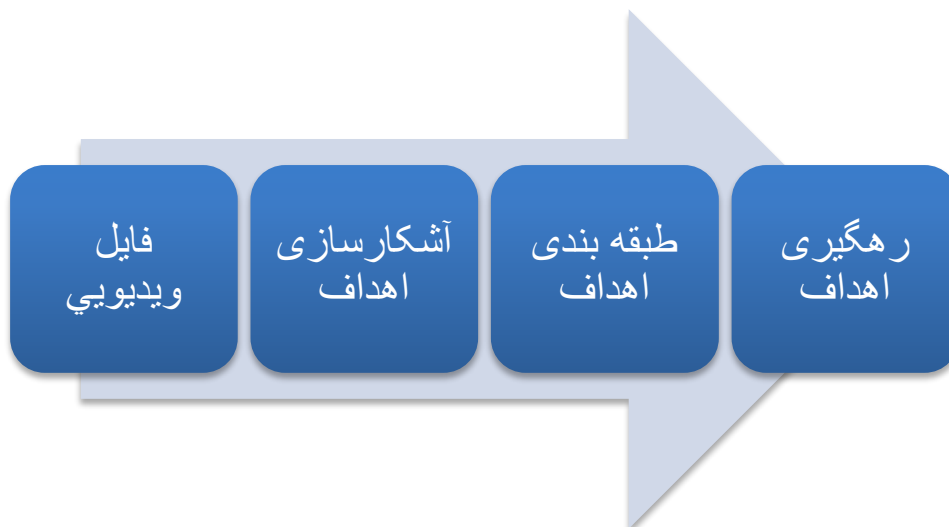
۲-۳- مراحل اصلی رهگیری هدف

فرآیند رهگیری دارای دو مرحله پیش پردازش اساسی است که نقش حیاتی را در رهگیری دارند که عبارت از آشکارسازی هدف و طبقه‌بندی هدف می‌باشد. هر چند در نگاه اول ممکن است اختلاف سه مرحله آشکارسازی هدف، جداسازی (طبقه‌بندی) هدف و رهگیری هدف مورد اغفال واقع شود، اما هر کدام از مراحل یاد شده به خودی خود زمینه مطالعاتی جداگانه‌ای دارد. همانطور که در شکل زیر مشاهده می‌شود اولین مرحله، آشکارسازی اهداف و اشیا موجود در فریم تصویر می‌باشد. در مرحله بعد باید بتوان اهداف را بسته به آن چه می‌خواهیم رهگیری کنیم، طبقه‌بندی نماییم و سپس طبقه مورد نظر رهگیری شود.

^۱ . Frame work

^۲ . clue

^۳ . non rigid



شکل ۲-۱ مراحل اصلی رهگیری هدف

۲-۳-۱- آشکارسازی هدف:

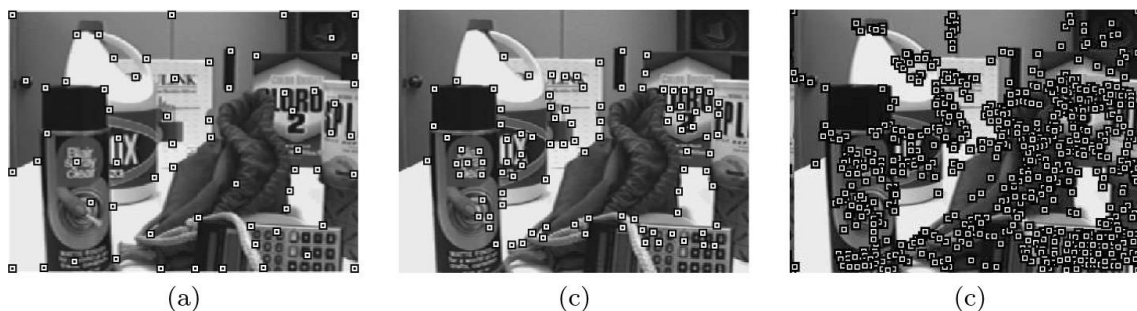
آشکارسازی هدف یک فناوری کامپیوتری است که با تشخیص نمونه‌های semantic از یک طبقه خاص نظیر انسان‌ها، ساختمان‌ها و یا خودروها در عکس و یا فیلم همراه است. آشکارسازی هدف می‌تواند با استفاده از برخی روش‌ها نظیر جداسازی قاب، جریان نور و یا تفکیک پس زمینه انجام شود. در بعضی از کاربردهای مربوط به دنبال کردن شیء هنگامی که برای اولین بار شیء وارد زاویه دید دوربین می‌شود، باید آشکارسازی شود و در بعضی از کاربردها نیز که چند شیء باید به طور همزمان دنبال شوند، در هر فریم آشکارسازی صورت می‌گیرد و پس از آن این وظیفه الگوریتم دنبال کردن شیء است که تعیین کند که هر یک از اشیاء آشکارسازی شده متناظر با کدام شیء در فریم قبلی می‌باشند. یک روش متداول برای آشکارسازی شیء متحرک در یک رشته ویدئویی استفاده از اطلاعات یک فریم می‌باشد. در برخی از روشها نیز اطلاعات زمانی^۱ که با در نظر گرفتن رشته ای از فریمها بدست می‌آید مورد استفاده قرار داده می‌شود. این اطلاعات زمانی، معمولاً به شکل تفاضل فریمها هستند که نواحی تغییر کننده

^۱ -temporal information

در فریمهای متوالی را برجسته می کند. در ادامه روشهای مختلف آشکارسازی را بصورت خلاصه بررسی می-کنیم.

۲-۳-۱-۱- آشکارسازی نقاط مهم

این روشها برای یافتن نقاط مهم که در توصیف بافت و انطباق تصاویر نقش موثری دارند، به کار می روند. از جمله این آشکارسازها میتوان به آشکارسازهای Harris^۱، KLT^۲، SIFT اشاره کرد [۲۴]. در شکل ۲-۲ تصویر یک صحنه و نقاط آشکارسازی شده در آن به سه روش مختلف آورده شده است.



شکل ۲-۲: آشکارسازی نقاط مهم، (a) The Harris, (b) The KLT, and (c) SIFT operators [۲۴]

۲-۳-۱-۲- تفریق پس زمینه

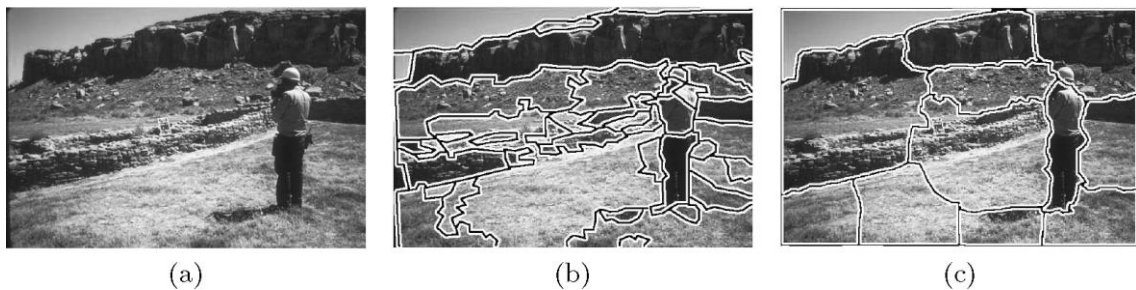
آشکارسازی شیء را می توان با تهیه مدلی برای پس زمینه و سپس یافتن انحراف ها از این مدل برای هر فریم ورودی انجام داد. هر تغییر مهمی در ناحیه ای از تصویر نسبت به مدل پس زمینه، دلالت بر وجود یک شیء متحرک دارد. پیکسلهایی که ناحیه تغییر کننده را تشکیل می دهند، برای پردازشهای بعدی علامت گذاری می شوند. معمولاً یک الگوریتم بخش متصل برای بدست آوردن نواحی متصل متناظر با اشیاء اعمال می شود [۲۴].

^۱ Kanade-Lucas-Tomasi

^۲ Scale Invariant Feature Transform

۲-۳-۱-۳- بخش بندی تصویر

هدف الگوریتمهای بخش بندی تصویر، قسمت بندی تصویر به ناحیه هایی است که از لحاظ ادراکی مشابهند (شکل (۲-۳)). هر الگوریتم بخش بندی دو مسئله را مشخص می کند اول معیاری برای یک بخش-بندی خوب و دوم روشی کارآمد برای قسمت بندی. از جمله روشهای بخش بندی می توان به روشهای بخش بندی با استفاده از روش انتقال میانگین و گراف کاتنها و کانتورهای فعال اشاره کرد. شکل ۲-۳ نمونه ۱ از بخش بندی تصویر با استفاده از روش های ذکر شده را نشان می دهد [۲۴].



شکل ۲-۳ : بخش بندی تصویر یا استفاده از انتقال میانگین (b) و گراف کاتنها (c) [۲۴]

۲-۳-۱-۴- یادگیری با سرپرست

آشکار سازی شیء متحرک را می توان با یادگیری نماهای مختلف شیء به کمک مجموعه ای از مثالها انجام داد. یادگیری نماهای مختلف شیء نیاز به ذخیره سازی مجموعه کامل الگوها را از بین برد. با استفاده از مجموعه ای از مثالهای محدود یادگیری، این روش قادر است نماهای نادیده شیء را بیاموزد. در زمینه آشکار سازی شیء مثالهای یادگیری از زوجهایی از ویژگی های شیء و کلاس شیء اختصاص داده شده به آنها ترکیب شده اند که هر دوی این کمیت ها به صورت دستی تعریف می شوند [۲۴].

انتخاب ویژگی ها نقش مهمی در کارایی الگوریتم طبقه بندی ایفا می کند. بنابراین مهم است که از ویژگی هایی استفاده شود که یک کلاس را از کلاس دیگر تمیز دهند. هنگامی که ویژگی های شیء انتخاب شده اند، می توان یادگیری ظاهرهای مختلف از یک شیء را با استفاده از یک روش یادگیری با سرپرست

انجام داد. شبکه های عصبی بوستینگ تطبیقی درختهای تصمیم و ماشین بردار پشتیبان از جمله روشهای کارآمد در این زمینه می باشند. این روشهای یادگیری، یک ابر رویه که یک کلاس شیء را از کلاسهای شیء دیگر جدا می کند را محاسبه می کنند.

۲-۳-۲- طبقه بندی اهداف

بعد از آشکارسازی اهداف در یک رشته ویدیویی، مرحله بعد تشخیص این اهداف و طبقه بندی آنها با توجه به نیازها است. طبقه بندی فرآیندی است که توسط آن اهدافی که در یک فریم آشکار شده است بر مبنای علاقه مندی طبقه بندی می شوند. شناسایی و طبقه بندی اهداف توسط پارامترهایی نظیر شکل، حرکت، رنگ و بافت انجام می شود و بر مبنای آنها تقسیم بندی انجام می شود. بسته به اینکه چه پارامترهایی برای طبقه بندی انتخاب شده باشند، روشهای مختلفی تعریف می شوند: طبقه بندی بر مبنای شکل، طبقه بندی بر مبنای حرکت، طبقه بندی بر مبنای رنگ و طبقه بندی بر مبنای بافت.

۲-۳-۱- طبقه بندی بر مبنای شکل:

این نوع طبقه بندی بر مبنای تحلیل شکل صورت می پذیرد. آنالیز شکل یک تحلیل خودکار از شکل های هندسی است که با مقایسه شکل اشیا با داده مرجع میزان شباهت بین آنها تشخیص داده می شود. در اکثر کاربردها نمایش مرز اشیا مورد استفاده قرار می گیرد هرچند نمایش های مبتنی بر حجم یا نقطه نیز کاربرد- های مخصوص بخود را دارند. یک نمایش ساده شده از شکل توصیف کننده شکل نامیده می شود. یک توصیف کننده شکل کامل شامل همه اطلاعات مورد نیاز برای باسازی شکل می باشد. توصیف کننده های شکل می توانند با توجه به کاربرد آنها متفاوت باشند از قبیل توصیف کننده های همسان، توصیف کننده های مبتنی بر گراف و کلاس های دیگر. همچنین شکل های می توانند بر اساس ساختار بخشی طبقه بندی شوند. اما بکارگیری ساختار بخشی با توجه به خاصیت غیرخطی بودن شکل ها فرآیند ساده ای نیست. ساختار

بخشی بصورت کلی در سه گروه طبقه‌بندی می‌شود: گروه اول بخش‌های سازنده شی را با توجه به تصاویر نمونه برداری شده مدل می‌کند. این فرآیند نیازمند داشتن یک دانش اولیه از بخش‌هاست. دو گروه دیگر بخش‌های سازنده را فقط از یک تصویر بدست می‌آورند. گروه دوم بخش‌های خاصی از یک شی در فریم‌های متفاوت با هم مقایسه می‌شوند تا معیار شباهت برای آنها در همه فریم‌ها بدست آید و در گروه سوم ساختار بخشی با توجه به ناحیه داخلی مرز شکل ایجاد می‌شود [۲۵].

۲-۲-۳-۲- طبقه‌بندی بر اساس حرکت:

طبقه‌بندی بر مبنای حرکت بر روی حالت تناوبی حرکت کار می‌کند. یک سیستم می‌تواند بگونه‌ای ساخته شود که چگونگی حرکت اشیا را در حین فرآیند دنبال کردن یاد بگیرد و بدینوسیله می‌تواند آن را بهتر طبقه‌بندی کند. در طبقه‌بندی مبتنی بر حرکت می‌بایست حرکت دو نوع از اشیا صلب و غیر صلب بصورت جداگانه مورد بررسی قرار گیرد. آن واضح است که فرآیند رهگیری اشیا صلب آسانتر است و برخلاف اشیا غیرصلب محدودیتی برای نشان دادن حالت تناوبی حرکت ندارند [۲۶].

۲-۲-۳-۲- طبقه‌بندی مبتنی بر رنگ:

رویکردهای مبتنی بر رنگ خصوصیات رنگ در هر تصویر را مورد مطالعه قرار می‌دهند. دو نوع خصوصیات اصلی رنگ که برای طبقه‌بندی بر مبنای رنگ مورد استفاده قرار می‌گیرند، توزیع قدرت طیف روشنایی و ویژگی انعکاس سطح اشیا است. اطلاعات رنگ معمولاً در فضای RGB نمایش داده می‌شود. مشکل فضای رنگ RGB این است که فضایی یکنواخت نیست بنابراین با توجه به نوع کاربرد و مسئله مورد نظر در فرآیند رهگیری استفاده از فضاهای رنگ دیگری از قبیل L^*a^*b و L^*u^*v بدلیل یکنواخت بودن می‌تواند نتیجه بهتری ارائه دهد. در همه فضاهای رنگ ذکر شده ویژگی برای تعریف یک شی کافی و موثر نیستند. از این روی در سال‌های اخیر توصیف‌گرهای رنگ به دو دسته توصیف کننده رنگ مبتنی بر هیستوگرام و توصیف-گرهای مبتنی بر SIFT تقسیم شده‌اند [۲۷].

۲-۳-۴- طبقه‌بندی بر مبنای بافت:

بافت یکی از ویژگی‌های ذاتی تمامی سطوح در تصویر است مانند زبری چوب، موجدار بودن پارچه، الگوی محصولات کشاورزی در زمین و غیره. بافت شامل اطلاعات مهمی درباره نظم ساختاری سطوح و ارتباط آنها با محیط اطراف است. از این روی ویژگی‌های بافتی تصاویر برای تمایز سطوح‌های مختلف با بافت‌های متفاوت بکار گرفته می‌شوند [۲۸]. توصیف‌کننده‌های بافت برای مشخص کردن میزان هم جنس بودن سطح و آشکارسازی ترکیب هیستوگرام ناحیه مورد نظر بکار گرفته می‌شوند. از جمله مهم‌ترین توصیف‌کننده‌های بافت می‌توان به توصیف‌کننده هم‌جنسی بافت، توصیف‌کننده مرورگر بافت و توصیف‌کننده هیستوگرام لبه اشاره کرد.

۲-۳-۳- رهگیری هدف

رهگیری هدف روشی برای دنبال کردن یک هدف در میان فریم‌های متوالی به منظور تعیین چگونگی جابجایی هدف نسبت به سایر اشیای موجود در تصویر است. متداول‌ترین روش اندازه‌گیری موقعیت مرکز هدف در مختصات (x,y) در فریم‌های متوالی است. رهگیری هدف می‌تواند به انواع مبتنی بر نقطه، مبتنی بر هسته (kernel) و یا مبتنی بر سایه نما^۱ تقسیم می‌شود که در ادامه توضیح داده می‌شوند.

۲-۳-۳-۱- روش‌های رهگیری مبتنی بر نقطه

در روش رهگیری نقطه‌ای اهداف به صورت نقاطی در نظر گرفته می‌شوند که با ارزیابی موقعیت و چگونگی حرکتشان رهگیری می‌شوند. این عمل می‌تواند با استفاده از الگوریتم‌های فیلتر کالمن، فیلتر ذره-ای و یا رهگیری فرضیات چندگانه (MHT) انجام شود.

^۱ . silhouette

۲-۳-۱-۱- کالمن فیلتر^۱

کالمن فیلتر روشی است که از مجموعه‌ای از اندازه‌گیری‌های مشاهده شده در طول زمان استفاده می‌کند. این اندازه‌گیری‌ها می‌تواند شامل نویز و یا سایر خطاها باشد. کالمن فیلتر تخمینی از متغیرها تولید می‌کند که می‌تواند بسیار دقیق‌تر از تخمینی باشد که بر مبنای یک اندازه‌گیری به تنهایی بدست آید. به عبارت دیگر، کالمن فیلتر به طور بازگشتی بر روی جریانی از ورودی‌های آغشته به نویز عمل می‌کند و می‌تواند به تخمین آماری بهینه‌ای از وضعیت سیستم مورد مطالعه دست یابد [۲۹].

دو مرحله اساسی در کالمن فیلتر وجود دارد. مرحله اول پیش‌بینی است که تخمینی از وضعیت فعلی متغیرها را به همراه عدم قطعیت‌هایشان تولید می‌کند. سپس خروجی اندازه‌گیری بعدی مشاهده شده و تخمین‌ها با استفاده از میانگین وزندهی شده به روزرسانی می‌شوند. در این شرایط، به تخمین‌هایی که قطعیت بیشتری دارند وزن بیشتری اختصاص می‌یابد. از آنجا که این روش بازگشتی می‌باشد، تنها مقدار فعلی، قبلی و نیز ماتریس عدم قطعیت برای محاسبه زمان واقعی کافی است.

۲-۳-۱-۲- فیلتر ذرات^۲

فیلتر ذرات مجموعه‌ای از الگوریتم تخمین چگالی است که با استفاده از بکارگیری مستقیم معادلات بازگشتی بایسین، چگالی فضای حالت را تخمین می‌زند. هیچ فرضی در رابطه با دینامیک فضای حالت و یا تابع چگالی ایجاد نمی‌شود، اما روش مناسبی برای تولید نمونه از تابع توزیع دلخواه می‌باشد. نمونه‌ها با مجموعه‌ای از ذرات نمایش داده می‌شوند و به هر ذره یک ضریب وزنی اختصاص داده می‌شود. وزن هر ذره نمایانگر احتمال انتخاب آن نمونه از تابع توزیع احتمال است. هنگامی که متغیرهای حالت به طور نرمال توزیع نشده باشند (گوسی باشند)، کالمن فیلتر تقزیم ضعیفی ارائه می‌کند در حالیکه فیلتر ذره‌ای به منظور

^۱ . Kalman Filtering

^۲ . Particle Filtering

غلبه بر این مشکل استفاده می‌شود. فیلتر ذره‌ای از کانتورها، ویژگی‌های رنگ و یا نگاشت بافت استفاده می‌کند [۳۰].

۲-۳-۳-۳-۲- رهگیری چند فرضی^۱

این روش با مجموعه‌ای از فرضیات از مرحله قبلی که فرضیات والد نامیده می‌شوند و مجموعه‌ای از اندازه‌گیری‌ها که از ابتدا تا کنون انجام شده است شروع به کار می‌کند. هر فرضیه، مجموعه اندازه‌گیری‌ها را به نحو متفاوتی به مسیرها اختصاص می‌دهد. به عبارت دیگر، هر فرضیه نمایانگر مجموعه متفاوتی از ضرایب است که اندازه‌گیری‌ها را به مسیرهای متفاوت نسبت می‌دهد. با در نظر گرفتن مجموعه اندازه‌گیری جدید و یکی از فرضیات قبلی، یک فرضیه جدید تولید می‌شود که به عبارت دیگر اندازه‌گیری‌ها را به گونه‌ای جدید به مسیرها اختصاص می‌دهد [۳۱].

۲-۳-۳-۲- روش‌های رهگیری مبتنی بر کرنل^۲

در این روش‌ها، یک هدف بر مبنای محاسبه حرکت یک بیضی و یا مستطیل در فریم‌های متوالی رهگیری می‌شود. حرکت می‌تواند جابجایی، چرخش و یا تبدیل‌های پیوسته باشد. مشکل این روش آن است که ممکن است بخشی از هدف خارج هسته قرار گیرد در حالیکه بخشی از زمینه در داخل آن قرار بگیرد. رهگیری ممکن است بر مبنای شکل هندسی، ویژگی‌ها و یا ظاهر هدف باشد.

۲-۳-۳-۱- تطبیق الگوی نمونه^۳

تطبیق الگو روشی است که در پردازش تصویر دیجیتال برای یافتن بخش‌های کوچکی از یک عکس و یا ویدیو که با یک تصویر از پیش تعیین شده تطبیق داشته باشد بکار می‌رود. در این روش یک شاخص عددی

^۱ . Multiple Hypothesis Tracking(MHT)

^۲ . Kernel

^۳ . Simple Template Matching

محاسبه می‌شود که نشان می‌دهد تصویر موجود در فریم تا چه میزان با الگوی مورد نظر مطابقت دارد و از این طریق نواحی مورد نظر از فریم شناسایی می‌شوند.

۲-۳-۳-۲- روش جابجایی میانگین^۱

الگوریتم جابجایی میانگین یک روش غیرپارامتریک تحلیل فضای ویژگی است که برای پیدا کردن مقادیر حداکثر تابع چگالی بکار می‌رود. این الگوریتم مبتنی بر تکرار و بر اساس پیش‌بینی مقادیر ویژگی‌های با استفاده از مقادیر قبلی می‌باشد. یکی از ساده‌ترین روش‌ها برای این منظور، محاسبه نقشه اعتماد^۲ تصویر جدید بر مبنای هیستوگرام رنگی تصویر قبلی می‌باشد. در نتیجه حداکثر مقدار نقشه اعتماد در نزدیکی موقعیت پیش‌بینی شده برای هدف رخ می‌دهد. بنابراین، نقشه اعتماد یک تابع چگالی احتمال است و مقدار حداکثر تابع می‌تواند با استفاده از روش جابجایی میانگین محاسبه شود [۳۱].

ناحیه مورد نظر توسط یک مسیر بسته^۳ از اولین فریم ویدیو انتخاب می‌شود. سپس، تابع چگالی احتمال فعلی با تابع چگالی احتمال فریم‌های بعدی مقایسه می‌شود. میزان شباهت بین فریم‌ها توسط ضریب باهاتاچاریا نشان داده می‌شود. همین فرآیند تا رسیدن به آخرین فریم ویدیو ادامه داده می‌شود [۳۲]. این روش معایب زیادی دارد. تنها یک هدف قابل رهگیری است. انتخاب ناحیه مورد نظر باید به صورت دستی انجام شود. این روش نمی‌تواند اهدافی با سرعت بالا جابجا شوند را رهگیری کند.

۲-۳-۳-۳- ماشین بردار پشتیبان^۴

این روش از الگوریتم‌هایی استفاده می‌کند که توسط مدل‌های یادگیری نظارت می‌شوند و الگوریتم‌های یادگیری به کار برده شده داده را تحلیل کرده و الگوها را بر اساس داده‌های تعلیم شده قبلی تشخیص می‌دهند. ماشین بردار پشتیبان اغلب برای طبقه‌بندی و رگرسیون استفاده می‌شوند. ایده اصلی ماشین بردار

^۱ . Mean Shift Method

^۲ . Confidence map

^۳ . Bounding box

^۴ . Support Vector Machin

پشتیبان تقسیم نقاط به دو مجموعه جدا از هم می‌باشد. اشیایی که قرار است رهگیری شوند در گروهی جدا از اشیایی که نیاز به رهگیری ندارند تقسیم‌بندی می‌شوند. این روش نیاز به راه‌اندازی و شروع به کار دستی^۱ داشته و بواسطه یادگیری، نیاز به بار محاسباتی بالایی دارد. این روش نمی‌تواند تصادم جزئی را مدیریت کند.

۲-۳-۳-۴- رهگیری مبتنی بر لایه^۲

این روش قابلیت رهگیری چند هدف را دارا می‌باشد. در این روش، هر لایه متناظر با یک شکل خاص است و بر مبنای آن هدف مورد نظر در آن لایه رهگیری می‌شود. از آنجا که وجود تعداد زیادی از این نوع لایه‌ها امکان‌پذیر است رهگیری چند هدف به صورت همزمان امکان‌پذیر است. هر لایه سه ویژگی مختلف دارد که مشخص می‌کند در آن لایه خاص چه هدفی باید رهگیری شود. هر لایه بر مبنای شدت نور یک هدف، حرکت و ظاهر آن طراحی می‌شود. این روش توانایی رهگیری چند هدف به طور همزمان را دارا می‌باشد. همچنین این روش توانایی مدیریت رهگیری در شرایطی که تصادم کلی نیز رخ می‌دهد را دارا می‌باشد [۳۲].

۲-۳-۳-۳- رهگیری سایه نما

این روش در زمان رهگیری اهداف با شکل‌های پیچیده که امکان نمایش آن‌ها با اشکال هندسی ساده امکان‌پذیر نیست استفاده می‌شود. اشکال پیچیده شامل انگشت، دست، شانه و از این قبیل می‌باشد. هدف این روش تولید یک مدل برای شی مورد نظر از فریم قبلی و مقایسه آن با ناحیه‌ای از فریم بعدی که شی در آن قرار دارد می‌باشد و به این ترتیب رهگیری انجام می‌شود.

۲-۳-۳-۱- رهگیری با استفاده از کانتور

^۱ . Manual

^۲ . Layer based tracking

در این روش کانتور مربوط به هدف از فریم قبلی گرفته شده و جهت محاسبه کانتور در فریم بعدی به طور مکرر پردازش می‌شود. در این روش نیاز است میزان مشخصی از کانتور فریم قبلی با کانتور هدف در فریم بعدی همپوشانی داشته باشد. تنها هنگامی رهگیری به درستی انجام می‌شود که این نیازمندی برآورده شود.

دو راهکار برای اجرای رهگیری با استفاده از کانتور وجود دارد. در دیدگاه اول، حرکت و شکل کانتور با استفاده از مفهوم فضای حالت مدل می‌شود. دیدگاه دوم ساده‌تر بوده و در آن از روش‌های مستقیم بهینه‌سازی نظیر گرادیان کاهشی (gradient descent) برای حداقل کردن انرژی کانتور استفاده می‌شود. مهم‌ترین مزیت این روش امکان رهگیری اهداف با شکل‌های غیرعادی می‌باشد.

۲-۳-۳-۳-۲- تطبیق شکل

روش تطبیق شکل مشابه با روش‌های طبقه‌بندی مبتنی بر شکل است. اما به جای طبقه‌بندی اهداف، از اشکال برای رهگیری شکل خاصی استفاده می‌شود. این روش مشابه تطبیق الگو نیز می‌باشد چون اشکال در یک پایگاه داده ذخیره می‌شوند و سپس شکل موجود در فریم با شکل موجود در پایگاه داده مقایسه شده و سپس رهگیری انجام می‌شود. در این روش یک هدف قابل رهگیری بوده و تصادم نیز توسط تبدیل Hough مدیریت می‌شود [۳۲].

فصل سوم

تحلیل تئوریک الگوریتم‌های فیلتر کالمن و فیلتر

ذره‌ای

۳-۱- مقدمه

دو روش مهم که امروزه برای دنبال کردن شیء در یک رشته ویدئویی مطرح هستند، فیلتر کالمن^۱ و دیگری فیلتر ذره ای^۲ [۳۳-۳۴] می باشند.

در این فصل به معرفی روش هایی که از راهکار فضای حالت برای مدل کردن سیستم استفاده می کنند خواهیم پرداخت. در ابتدا فیلتر کالمن^۳ [۳۵] را مطرح می کنیم و با ارائه مثالی در این زمینه مفاهیمی مربوط به فضای حالت و چگونگی تعریف یک مسئله دنبال کردن شیء به صورت یک مسئله تعیین حالت و همچنین چگونگی حل مسئله در دو گام پیش بینی و بروز رسانی به روش بازگشتی را مطرح می کنیم. برای استفاده از فیلتر کالمن در یک مسئله تخمین حالت نیازمند در نظر گرفتن چند شرط محدود کننده هستیم، بدون در نظر گرفتن این شروط و در حالت کلی، می توان برای تخمین حالت سیستم از فیلتر ذره ای استفاده کرد. از این روی در ادامه به بررسی فیلتر ذره ای با چند معیار اندازه گیری شباهت بین فریم های مختلف می پردازیم.

۳-۲- فیلتر کالمن

در بسیاری موارد نیاز به تخمین حالت یک سیستم داریم در حالی که حالت سیستم در حال تغییر است و تنها از سیستم سنجش هایی نویزی در اختیار داریم. برای مثال در یک سیستم دنبال کردن شیء بردار حالت را می توان در لحظه K ام به صورت زیر در نظر گرفت:

$$X_k = \begin{bmatrix} x_k \\ y_k \\ v_{xk} \\ v_{yk} \end{bmatrix} \quad (1-3)$$

^۱ - Mean Shift

^۲ - Particle Filter

^۳ - Kalman Filter

که در آن (x_k, y_k) موقعیت شیء در صفحه را مشخص می کند، v_{xk} سرعت شیء در جهت x و v_{yk} سرعت شیء در جهت y می باشد.

و سنجشهای در دسترس با بردار Z_k مشخص می شوند:

$$Z_k = \begin{bmatrix} xm_k \\ ym_k \end{bmatrix} \quad (2-3)$$

عناصر این بردار موقعیت شیء در فریم جاری می باشند که توسط یک روش آشکار سازی شیء بدست آمده است.

در فیلتر کالمن همچنین فرض می شود که برای تعیین حالت سیستم در لحظه $K+1$ ام دو معادله زیر (معادله^۱ پروسه و معادله سنجش) را در اختیار داریم:

$$X_{k+1} = F_{k+1,k} X_k + w_k \quad (3-3)$$

که معادله پروسه نامیده می شود. در این معادله $F_{k+1,k}$ ماتریس انتقال^۲ و w_k نویز سیستم می باشد که در روش فیلتر کالمن این نویز دارای توزیع گوسی با میانگین صفر و کواریانس Q_k فرض می شود. معادله دوم معادله سنجش^۳ است که رابطه بین موقعیت واقعی شیء و موقعیت اندازه گیری شده را در حضور عدم قطعیت ها مشخص می کند. این رابطه به صورت زیر می باشد:

$$Z_k = H_k X_k + V_k \quad (4-3)$$

در این رابطه Z_k مشاهده ما در زمان K ، H_k ماتریس سنجش^۴ و V_k نویز اندازه گیری می باشد که دارای توزیع گوسی با میانگین صفر و ماتریس کواریانس R_k می باشد. برای اینکه در این مسئله بتوانیم از فیلتر کالمن استفاده کنیم باید تابع چگالی احتمال اولیه $p(X_0|Z_0)$ به صورت گوسی باشد همچنین نویزهای

^۱- Process equation

^۲- Transition matrix

^۳- Measurement equation

^۴- Measurement matrix

سیستم و اندازه گیری با توزیع گوسی فرض شوند و توابع پیاده سازی شده با ماتریس های H_k , $F_{k+1,k}$ توابعی خطی باشند.

برای مثال معادله روند و معادله سنجش می توانند برای مسئله دنبال کردن شیء به صورت زیر پیشنهاد شوند.

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ v_{xk+1} \\ v_{yk+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ v_{xk} \\ v_{yk} \end{bmatrix} + W_k \quad (5-3)$$

$$\begin{bmatrix} xm_k \\ ym_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ v_{xk} \\ v_{yk} \end{bmatrix} + V_k \quad (6-3)$$

روش کار فیلتر کالمن را می توان در دو مرحله بیان کرد [۳۵] مرحله اول پیش بینی و مرحله دوم به روز

رسانی. در مرحله پیش بینی با استفاده از حالت قبلی حالت فعلی تخمین زده می شود این تخمین را با $\hat{X}_K^-$ تصحیح می گردد.

$$\hat{x}_k^- = F_{k,k-1} \hat{x}_{k-1} \quad (7-3)$$

و در مرحله به روز کردن با توجه به موقعیت اندازه گیری شده، $\hat{X}_K^-$ تصحیح می گردد.

دو گام فوق را می توان با رابطه زیر نمایش داد

$$\hat{x}_k = G_k^{(1)} \hat{x}_k^- + G_k z_k \quad (8-3)$$

که در آن ماتریسهای $G_k, G_k^{(1)}$ باید تعیین شوند و $\hat{x}_k$ تخمین نهایی فیلتر کالمن می باشد.

با استفاده از روابط ریاضی می توان نشان داد [۳۶] :

$$G_k^{(1)} = I - G_k H_k \quad (9-3)$$

بنابراین رابطه (۳-۸) به صورت زیر قابل بیان می باشد:

$$\hat{x}_k = \hat{x}_k^- + G_k (z_k - H_k \hat{x}_k^-) \quad (10-3)$$

که در آن ماتریس G_k بهره کالمن^۱ نامیده می شود. می توان نشان داد [۳۶] :

$$G_k = P_k^- H_k^T [H_k P_k^- H_k^T + R_k]^{-1} \quad (11-3)$$

در این P_k^- ماتریس کواریانس پیشین می باشد که آن را با رابطه زیر معرفی می کنیم:

$$P_k^- = E \left[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T \right] \quad (12-3)$$

رابطه (۱۱-۳) بهره کالمن را برای حل معادله (۳-۱۰) با فرض داشتن حالت قبلی که تنها فاکتور مجهول

می باشد، تعیین می کند اما مشکل در اینجا است که بهره کالمن خود تابعی از ماتریس کواریانس پیشین P_k^-

می باشد که مجهول است. با توجه به اینکه x_k را در دسترس نداریم نمی توانیم P_k^- را مستقیماً محاسبه

کنیم لذا برای تکمیل عملکرد فیلتر از یک روش بازگشتی استفاده می کنیم. این روش شامل دو مرحله می

باشد:

(۱) با توجه به ماتریس پسین مربوط به مرحله قبل، P_{k-1} ، ماتریس کواریانس پیشین P_k^- را به روز می

کنیم:

$$P_k^- = F_{k,k-1} P_{k-1} F_{k,k-1}^T + Q_{k-1} \quad (13-3)$$

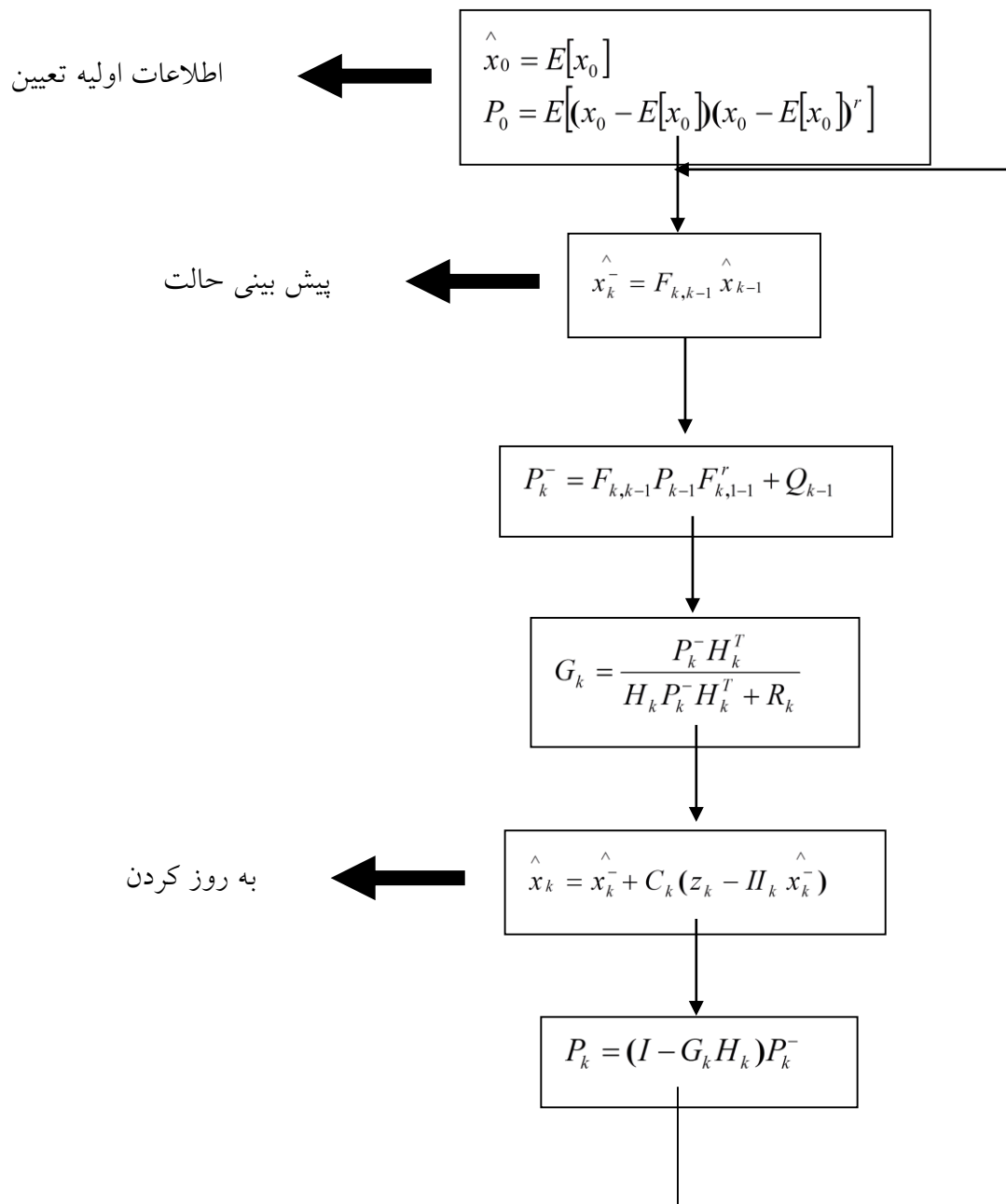
(۲) با استفاده از ماتریس کواریانس پیشین P_k^- ، ماتریس کواریانس پسین را با رابطه زیر بدست می

آوریم:

$$P_k = (I - G_k H_k) P_k^- \quad (14-3)$$

روند نمای زیر پروسه انجام شده در فیلتر کالمن را نشان می دهد.

^۱ - Kalman gain



شکل ۳-۱: روند نمای فیلتر کالمن

۳-۳- فیلتر ذره ای

همانطور که در بخش قبل گفته شد برای استفاده از فیلتر کالمن باید چند شرط محدود کننده را در نظر بگیریم به این صورت که تابع چگالی احتمال اولیه $p(x_0|z_0)$ باید به صورت گوسی باشد همچنین باید نویزهای سیستم و اندازه گیری با توزیع گوسی فرض شوند و توابع پیاده سازی شده با ماتریسهای $H_k, F_{k+1,k}$ توابعی خطی باشند. برای سیستمهایی که در آنها این شروط برآورده نمی شوند، فیلتر کالمن قابل استفاده نمی باشد و به جای آن می توان از فیلتر ذره ای [۳۳-۳۴] استفاده کرد. فیلتر ذره ای تکنیکی بر پایه تخمین بیزین بازگشتی برای حالت مجهول سیستم می باشد. از جمله دلایل محبوبیت فیلتر ذره ای می توان به این موارد اشاره کرد: توانایی دنبال کردن چند فرض، توانایی کار کردن با مدل های غیر خطی و امکان پیاده سازی آسان و کارآمد.

در اینجا نیز مشابه قسمت قبل فرض می کنیم معادله پروسه و معادله سنجش را در اختیار داریم این معادلات را این بار به صورت زیر مطرح می کنیم.

معادله پروسه:

$$X_k = f_k(X_{k-1}, V_k) \quad (15-3)$$

و معادله سنجش:

$$Z_k = h_k(x_k, n_k) \quad (16-3)$$

که در این روابط f_k, h_k توابع غیرخطی و متغیر با زمان می باشند و V_k, n_k به ترتیب عدم قطعیت سیستم و نویز اندازه گیری را مدل می کنند. از دیدگاه بیزین مسئله دنبال کردن را می توان به صورت پیدا کردن تابع چگالی احتمال $p(x_k|z_{1:k})$ مطرح کرد. در این رابطه فرض می شود که تابع احتمال مربوط به حالت اولیه سیستم در اختیار باشد:

$$p(x_0|z_0) = p(x_0) \quad (17-3)$$

سپس دو گام پیش بینی و به روز کردن برای بدست آوردن تابع چگالی احتمال $p(x_k|z_{1:k})$ به کار

گرفته می شود فرض کنید $p(x_{k-1}|z_{1:k-1})$ در زمان $k-1$ در دسترس است. در این مرحله پیش بینی با استفاده از معادله حالت به شکل زیر صورت می گیرد:

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1}) p(x_{k-1}|z_{1:k-1}) dx_{k-1} \quad (18-3)$$

با داشتن $p(x_k|z_{1:k-1})$ و با استفاده از قاعده بیز داریم:

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k) p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})} \quad (19-3)$$

دو رابطه فوق اساس راه حل بیزین بهینه را شکل می دهند. این انتشار چگالی پسین در حالت کلی قابل تعیین به صورت تحلیلی نیست. در اینجا می توان روش فیلتر ذره ای را برای تخمین حالت سیستم به کار برد.

در روش فیلتر ذره ای فرض می شود در لحظه K تعداد N_s نمونه (ذره) با حالت های $\{S_k^{(n)}\}_{n=1,2,\dots,N_s}$ داریم و دنباله $\{\pi_k^{(n)}\}_{n=1,2,\dots,N_s}$ را که وزن مربوط به هر ذره را تعیین می کند نیز در اختیار داریم. سپس چگالی پسین در نقاط حضور ذره ها به صورت زیر فرض می شود:

$$p(x_k|z_{1:k}) \approx \sum_{n=1}^{N_s} \pi_k^{(n)} \delta(x_k - s_k^{(n)}) \quad (20-3)$$

که در آن وزن $\pi_k^{(n)}$ عددی متناسب با $p(z_k|s_k^{(n)})$ در نظر گرفته می شود.

در اینجا الگوریتم موسوم به انتشار چگالی شرطی^۱ [۳۴] که یکی از مشهورترین روش های پیاده سازی فیلتر ذره ای در زمینه بینایی ماشین می باشد را مطرح می کنیم:

^۱ - Condensation

از مجموعه نمونه قدیمی $\{s_{k-1}^{(n)}, \pi_{k-1}^{(n)}, c_{k-1}^{(n)}, n=1 \dots N\}$ در زمان $t-1$ مجموعه نمونه جدید

$\{s_k^{(n)}, \pi_k^{(n)}, c_k^{(n)}, n=1 \dots N\}$ در زمان t را تولید کن.

n امین نمونه از N نمونه را به شکل زیر تولید کن:

۱- یک نمونه $s_k^{(n)}$ را به صورت زیر انتخاب کن:

(الف) یک عدد تصادفی $r \in [0,1]$ با توزیع یکنواخت را تولید کن.

(ب) کوچکترین j را که به ازای $c_{k-1}^{(j)} \geq r$ را پیدا کن.

(ج) قرار بده $s_k^{(n)} = s_{k-1}^{(j)}$.

۲- پیش بینی را برای انتخاب هر $s_k^{(n)}$ با نمونه گیری از $p(x_k | x_{k-1} = s_k^{(n)})$ انجام بده.

۳- وزن موقعیت جدید را بر حسب مشخصه های مشاهده شده z_t اندازه گیری کن:

$$\pi_k^{(n)} = p(z_k | x_k = s_k^{(n)})$$

سپس به گونه ای نرمالیزه کن که $\sum_n \pi_k^{(n)} = 1$. و در کنار احتمال تجمعی به صورت $\{s_k^{(n)}, \pi_k^{(n)}, c_k^{(n)}\}$

قرار بده که در آن

$$\begin{aligned} c_k^{(0)} &= 0, \\ c_k^{(n)} &= c_k^{(n-1)} + \pi_k^{(n)} \quad (n=1, \dots, N) \end{aligned}$$

هنگامیکه N نمونه تولید شدند: در صورت تمایل، ممان موقعیت دنبال شده را در زمان t با صورت زیر

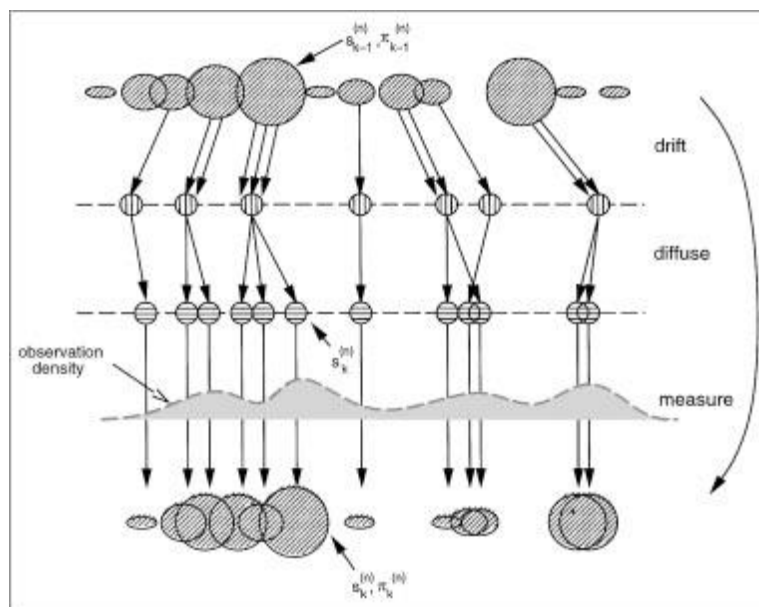
محاسبه کن:

$$\varepsilon \left[f(x_k) = \sum_{n=1}^N \pi_k^{(n)} f(s_k^{(n)}) \right]$$

برای مثال موقعیت میانگین با $f(x)=x$ بدست می آید.

الگوریتم ۱: الگوریتم انتشار چگالی شرطی

این عملکرد را می توان با استفاده از شکل ۳-۲ توضیح داد. در این شکل در سطر بالایی، هر بیضی بیانگر یک ذره در فیلتر مورد بحث و مساحت هر بیضی متناسب با مشاهده (سنجش) قبلی $p(z_{k-1}|s_{k-1}^{(n)})$ می باشد. در مرحله بعد نمونه برداری مجدد صورت می گیرد. این نمونه برداری مجدد به گونه ای انجام می شود که امکان واقع شدن نمونه جدید در محل نمونه ای که وزن بیشتری دارد، بیشتر خواهد بود. پس از آن معادله پروسه بر روی نمونه های جدید اعمال می شود (سطر سوم و چهارم در شکل ۳-۲). پس از آن هر نمونه با توجه به مشاهده جدید (z_k) وزن دهی می شود (سطر آخر در شکل) و در نهایت در محل هر نمونه احتمال پسین با توجه به معادله (۳-۲۰) قابل محاسبه خواهد بود.



شکل ۳-۲ عملکرد الگوریتم انتشار چگالی شرطی

فیلتر ذره ای مسئله تخمین حالت غیرگوسی و غیرخطی را حل می کند و می تواند با توابع احتمال چند مدی کار کند و توانایی بازیابی در صورت گم کردن مسیر شیء را دارا می باشند. لیکن برای پوشش مناسب فضای حالت معمولاً نیاز به تعداد زیادی ذره می باشد. این تعداد با افزایش ابعاد فضای حالت به شدت افزایش می یابد و باعث افزایش هزینه محاسباتی می گردد.

۳-۴- علل نیاز به رویکرد نوین

دنبال کردن عابران پیاده در صحنه‌های واقعی یک مولفه بحرانی برای بسیاری از کاربردهاست، از قبیل سیستم‌های نظارت ویدیویی و سیستم‌های دستیار راننده. شناسایی و طبقه‌بندی دقیق عابران در بهبود فرآیند دنبال کردن عابران موثر است. اخیراً الگوریتم‌های دنبال کردن عابران پیاده توانسته اند کارایی خوبی را در صحنه های کم ترد بدست آورند اما کارایی آنها در صحنه های شلوغ بشدت کاهش می یابد بطوریکه ابهام در شکل ظاهری به سبب تصادم‌های زیاد ارتباط فضایی پیچیده ای را در بین عابران مختلف ایجاد می-کند.

غالب روش‌های شناسایی عابران پیاده بدین صورت کار می‌کند: دریافت تصویر، حرکت دادن پنجره با چند اندازه بر روی کل تصویر و بکارگیری یک کلاسیفایر باینری برای مشخص کردن اینکه آیا عابر پیاده در پنجره وجود دارد یا خیر. معمولاً یک عابر پیاده بوسیله پنجره های مختلفی مرزبندی می‌شود، $non\ maxima$ suppression (NMS) اغلب برای تطبیق رویهم افتادگی های آشکارسازی شده با توجه به معیار رویهم افتادگی استفاده می‌شود. حداقل دو ضعف برای این روشها در فضاهای شلوغ وجود دارد. اولاً شکل ظاهری عابران بدلیل تصادم‌های شدید بسیار تغییر می‌کند بنابراین برای یک طبقه‌بندی کننده باینری شناسایی تمامی عابران پیاده بدرستی بسیار مشکل خواهد بود. دوماً در صحنه های شلوغ معمولاً چندین عابر پیاده بسیار نزدیک بهم هستند و از این روی بکارگیری یک سطح استانه ثابت نمی تواند بدرستی عابران را شناسایی کند.

از طرف دیگر الگوریتم‌های بررسی شده نیز در فضای حالت مانند کالمن و فیلتر ذره‌ای توانایی مقابله با تصادم‌های طولانی مدت و حل مسئله‌ای با پیچیدگی بالا و شباهت ظاهری زیاد اهداف را ندارند. لذا برای دنبال کردن عابران پیاده در یک فضای شلوغ، پر تصادم با اهدافی که بسیار شبیه بهم هستند نیاز به روش-های جدید در این حوزه احساس می‌شود.

فصل چهارم

مدل کردن مسئله رهگیری چند هدف با ابرگراف

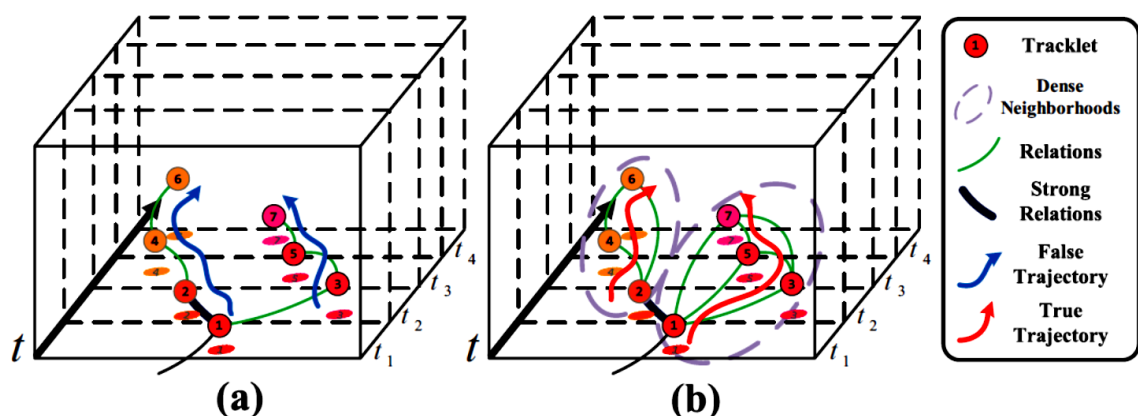
۴-۱- مقدمه

دنبال کردن چند هدف همزمان یکی از مسائل جالب اما چالش برانگیز در بینایی ماشین می‌باشد. اغلب روشهای قبلی انتساب داده بر مبنای روشهای بررسی ارتباط بین شباهت های شکل ظاهری و الگوی حرکت استوار بودند. از این روی فرآیند دنبال کردن در صحنه های با تصادم زیاد و حرکت اهداف شبیه بهم در فاصله های بسیار نزدیک از هم در این روشها با مشکل مواجه خواهد شد. در الگوریتم پیشنهادی یک روش انتساب داده جدید بصورت سلسله مراتبی بر مبنای ابر گراف بدون جهت پیشنهاد شده است، که مسئله دنبال کردن یک هدف بعنوان یک مسئله جستجوی سلسله مراتبی همسایه های متراکم (نزدیک بهم) بر روی گراف شباهت ساخته شده بصورت پویا فرمول بندی می کند [۳۷]. رابطه بین آشکارسازی های مختلف در سرتاسر حوزه زمان-مکان در یک روش مرتبه بالا مورد بررسی قرار گرفته است که دنبال کننده را قادر می سازد در زمان هایی که اهداف نزدیک بهم هستند یا ظاهری شبیه بهم دارند، فرآیند رهگیری را با دقت بالا دنبال کند. از این روی طراحی سلسله مراتبی فرآیند بهینه سازی دنبال کننده را در مواجهه با صحنه های طولانی مدت تصادم و برخورد قدرتمند می کند.

دنبال کردن همزمان چندین هدف در عین جذاب بودن مسئله بسیار مشکلی است. در ادبیات موضوعی روشهای بسیار زیادی پیشنهاد شده است، اما کارایی آنها در عمل راضی کننده نیست. همانگونه که در شکل ۱ نشان داده شده است اکثر روشهای قبلی بجای متمرکز شدن بر روی چندین مسیرک در کل حوزه زمانی یک ویدیو، بر روی رابطه دودو مسیرکها (بعنوان مثال شباهت ظاهری و الگوی حرکت) در حوزه زمانی-مکانی محدود متمرکز شده اند. هنگامی که اهداف در نزدیکی هم با ظاهری مشابه و الگوی حرکتی شبیه بهم حرکت می کنند، همانگونه که بوسیله دایره در شکل ۴-۱ نمایش داده شده است، تشخیص هویت اهداف در روش های مورد مطالعه در تحقیقات قبل بخوبی صورت نمی گیرد [۳۸، ۳۹، ۴۰، ۴۱، ۴۲، ۴۳]. در [۴۴] روشی بر مبنای مینیمم کردن گراف های کلیک توسعه داده شده است که رابطه بین آشکارسازی های مختلف را در طول حوزه زمان بررسی می کند. با این وجود کنترل حرکت های غیرخطی اهداف در صحنه های شلوغ

بخصوص در هنگام رخ دادن تصادم مشکل است و بطور کلی منجر به تولید نودهای فرضی غیر واقعی در طی فرآیند بهینه سازی می گردد.

در این پایان نامه یک ابرگراف سلسله مراتبی غیر جهت دار (H^T) پیشنهاد شده است که مسئله دنبال کردن را بعنوان یک مسئله جستجوی چند همسایگی متراکم بر روی ساختار ابر گراف غیر جهت دار فرمول بندی می کند، همانگونه که در شکل ۴-۱ ب توصیف شده است.



شکل ۴-۱: (a) نشان می دهد که روش های قبلی مسیر حرکت اهداف را در هنگام داشتن ظاهری شبیه بهم و حرکت در نزدیکی بدرستی نمی توانند تشخیص دهند. (b) نحوه عملکرد الگوریتم پیشنهادی را بر اساس تحلیل سلسله مراتبی ابر گراف نشان می دهد [۳۷].

متفاوت از روش های قبل ما رابطه بین آشکارسازی های مختلف را در طول حوزه زمان بررسی کردیم. ضمناً یک استراتژی محلی به کل برای طبقه بندی شناسایی ها بصورت سلسله مراتبی بکار گرفته شده است که سبب کاهش بسیار زیاد پیچیدگی های محاسباتی در جستجوی همسایگی های متراکم در هنگام مواجهه با تغییرات ناگهانی ظاهری و حرکتی اهداف می شود.

نخست فایل ویدیوی به چند قسمت در حوزه زمان تقسیم می شود و همسایگی های متراکم در هر قسمت جستجو می شوند تا چندین مسیرک طولانی تر ساخته شوند. سپس قسمت های نزدیک بهم برای ساختن تقسیم بندی جدید جهت بکارگیری در لایه بعد باهم یکپارچه می شوند. این دو مرحله تکرار می شوند تا زمانی که فقط یک قسمت در لایه باقی بماند یعنی کل ویدیو مورد نظر و جستجوی همسایگی های متراکم

در این قسمت انجام می‌شود مسیر نهایی حرکت اهداف بدست آید. بخش‌های کلی روش پیشنهادی بصورت مختصر شامل موارد زیر می‌شود [۳۷]:

- ۱- دنبال کردن چند هدف در ابتدا بعنوان مسئله جستجوی همسایگی‌های متراکم بر روی ابرگراف وابستگی غیر جهت دار مسیرک‌های ساخته شده بصورت سلسله مراتبی مدل می‌شود.
- ۲- ویژگی‌های حرکت و اطلاعات ظاهری اهداف بوسیله بررسی رابطه مرتبه بالا مسیرک‌های متعدد در گراف ساخته شده بصورت کامل استخراج می‌شوند
- ۳- آزمایشات دنبال کردن روی تعدادی از فریم‌های چالش برانگیز عمومی نشان داد که روش پیشنهادی دارای عملکرد بسیار خوبی در هنگام مواجهه با تصادم‌های طولانی و چالش شباهت ظاهری در صحنه‌های شلوغ می‌باشد.

۴-۲- مروری بر تئوری ابر گراف

تئوری گراف یکی از مهمترین، ساده‌ترین و متداولترین مفاهیم مطرح‌شده در ریاضی برای مدلسازی و بیان مسائل مختلف می‌باشد. یک گراف $G(V,E)$ شامل تعداد دلخواهی راس^۱ (نود، گره، ...) و یال^۲ (لبه، کمان، لینک، ...) است. رئوس یک گراف بیانگر موجودیتها (شهرها، هنرپیشگان، انسانها، صفحات اینترنتی، اتمهای عناصر، ...) بوده و یالهای یک گراف، برقراری ارتباط میان این موجودیتها (راه ارتباطی، فیلم مشترک، رابطه دوستی، ارجاعات اینترنتی، وجود واکنش شیمیایی، ...) را نشان می‌دهد.

هر یال e مابین دو راس v_i و v_j از رئوس گراف قرار دارد که به کمک زوج نامرتب $e = (v_i, v_j)$ نشان داده می‌شود. مطابق تئوری گراف، خاصیت‌های متنوعی را می‌توان برای یک یال قائل شد که این امر باعث افزایش انعطاف و کارایی گراف مربوطه می‌شود. تعدادی از این خواص عبارتند از [۳۰]:

^۱ Vertex (V)

^۲ Edge (E)

- درگراف جهندار:
تمایز میان رئوس یک یال (بیان هر یال بصورت زوج مرتب از راس مبدا به راس مقصد)
- حلقه در گراف چندگانه:
یکسان بودن راس مبدا و راس مقصد برای یک یال: $e = (v_i, v_i)$
- یال دوبله در گراف چندگانه:
یکسان بودن رئوس مبدا و مقصد در دو یا چند یال: $e_1 = (v_i, v_j), e_2 = (v_i, v_j), \dots$
- در گراف برچسبی:^۲
اختصاص برچسب به یال (بیانگر نوع ارتباطات موجود میان رئوس)
- در گراف وزن دار:
اختصاص عددی به عنوان هزینه هر یال (مقایسه میان هزینه مسیرهای مختلف در یک گراف)

۴-۲-۱- برقراری ارتباط میان دو نوع متفاوت از موجودیت‌ها

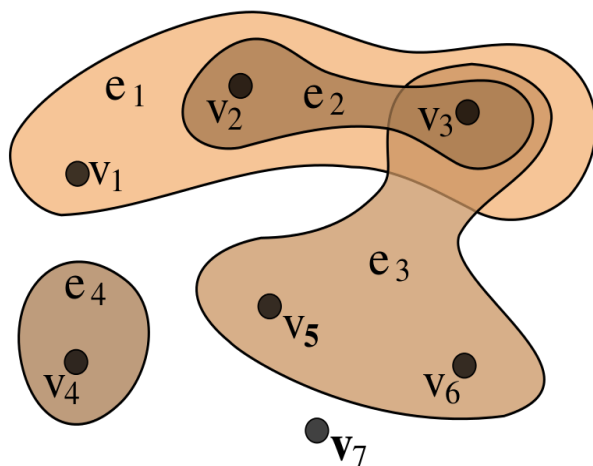
ابرگراف نیز یکی از تعمیم‌های مفهوم گراف است که در آن، یک یال می‌تواند تعداد دلخواهی از رئوس را به یکدیگر متصل نماید. بدین معنا که در ابرگراف، یک یال معادل مجموعه‌ای دلخواه از رئوس است در حالیکه هر یال در گرافهای معمولی معادل با یک زوج از رئوس بیان می‌شود.

تعریف ابرگراف به فرم $H = (X, E)$ است که در آن: X بیانگر مجموعه رئوس بوده و E یک مجموعه از زیرمجموعه‌های ناتهی از X است که ابريال^۳ نام دارند. در شکل زیر مثالی از یک ابرگراف و نمایش ریاضی آن دیده می‌شود:

^۱ Multigraph

^۲ Labeled Graph

^۳ Hyperedge



$$H = (X, E)$$

$$X = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$$

$$E = \{e_1, e_2, e_3, e_4\}$$

$$E = \left\{ \{v_1, v_2, v_3\}, \{v_2, v_3\}, \{v_3, v_5, v_6\}, \{v_4\} \right\}$$

شکل ۴-۲: ساختار کلی ابر گراف

در یک ابرگراف k -uniform، هر ابريال دقیقاً دارای k عدد از رئوس است. بدیهی است که برای حالت

$k = 2$ تعریف متداول یک گراف کلاسیک نتیجه می‌شود.

بسیاری از تعاریف و خواص موجود در یک زیرگراف به سادگی با تعمیم آن تعریف یا خاصیت از مفهوم

گراف بدست می‌آید. بعنوان مثال می‌توان به موارد زیر اشاره کرد:

- زیرابرگراف^۱: (تعمیم مفهوم زیرگراف)
- ابرگراف دوگان^۲: (تعمیم مفهوم گراف دوگان در تبدیل راسها و یالها به یکدیگر)
- ابرگراف همبند^۳: (تعمیم تعریف یک گراف همبند)
- ابرگراف دوقسمتی: (تعمیم مفهوم گراف دو قسمتی در افزاز راسهای گراف به دو مجموعه متفاوت)
- ابرگراف جهتدار: (تعمیم مفهوم گراف جهتدار: هر ابريال تعدادی راس مبدا را به تعدادی راس مقصد متصل می‌کند).

دو مثال از ابرگراف جهتدار در شکلهای زیر دیده می‌شود. ابريال e_1 در شکل سمت چپ، راس v_1 یا v_2 را

به راس v_3 یا v_4 متصل می‌کند (دو راس مبدا و دو راس مقصد). در همین شکل ابريال e_3 اتصال راس v_7 را

به v_3 یا v_4 برقرار می‌نماید (یک راس مبدا و دو راس مقصد). در شکل سمت راست، ابريال e برقراری اتصال

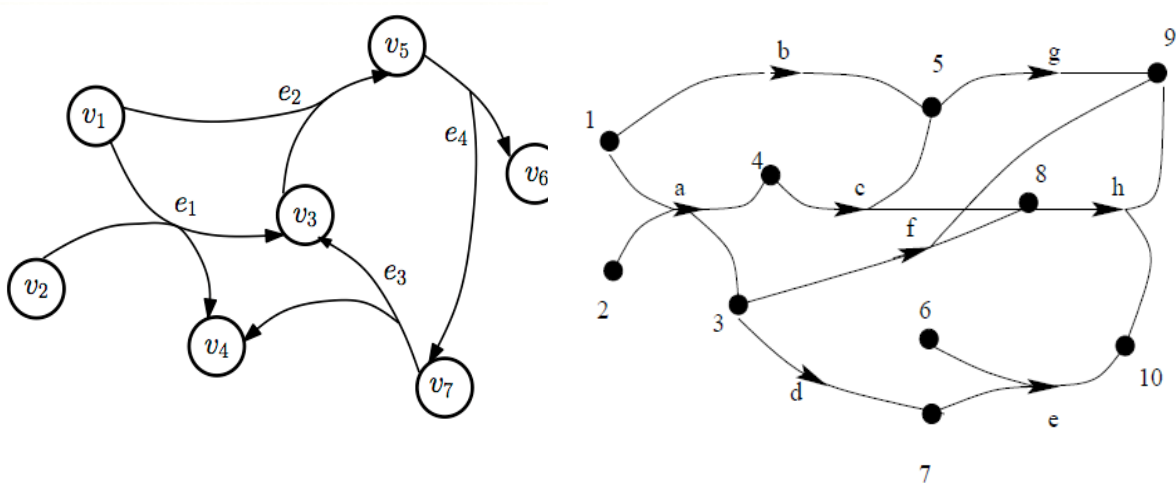
میان راس ۶ یا راس ۷ با راس ۱۰ را انجام می‌دهد (دو راس مبدا و یک راس مقصد) و یالهای b و g و

ارتباط میان یک راس مبدا با یک راس مقصد را برقرار می‌کنند.

^۱ Sub Hypergraph

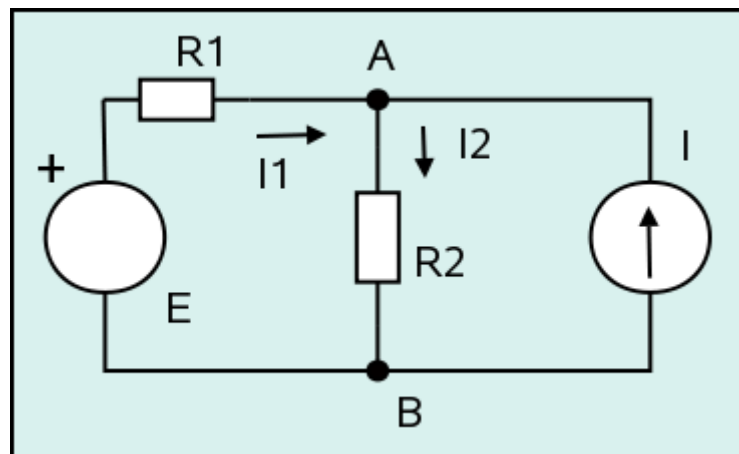
^۲ Dual Hypergraph

^۳ Connected Hypergraph



شکل ۴-۳: ابر گراف‌های جهت‌دار [۳۰]

البته مفاهیم مهمی مانند سیکل، مسیر، قطر، شعاع، قابلیت عبور، ... از جزییات و فرمولاسیون محاسباتی پیچیده‌تری نسبت به گراف‌های عادی برخوردار هستند. تعاریف متعددی ارتباط میان یک ابرگراف با یک (یا چند) گراف متناظر با آن را بیان می‌کنند. مثلاً مفهوم ۲-section یا clique برای ابرگراف، معادل با گرافی تعریف می‌شود که دارای رئوس یکسان با ابرگراف بوده و هر کدام از یال‌های رسم‌شده مابین دو راس از آن، در ابرگراف نیز وجود داشته باشد.



شکل ۴-۴: مدل کردن مدار با هایپر گراف

مفهوم ابرگراف کارایی بسیاری در نگاشت مسائل حوزه‌های مختلف به تئوری گراف و حل آنها دارد. مثلاً

مدل کردن مدار شکل مقابل به یک ابرگراف بدین صورت است که هریک از عناصر مدار را یک راس از

ابرگراف در نظر می‌گیریم (۴ راس شامل: E, I, R_1, R_2)، بدین ترتیب ابرگراف معادل این مدار دارای سه

ابرپال خواهد بود. مدل ابرگراف این مدار از مدل گرافی آن کاراتر و مفیدتر می‌باشد.

$$\text{Hyperedges} = \{\{R_1, R_2, I\}, \{E, R_2, I\}, \{R_1, E\}\}$$

کاربردهای فراوانی برای ابرگراف در کار با پایگاههای داده، سیستمهای خبره، مسیریابی در شبکههای

پویا، برنامه‌سازی موازی، داده‌کاوی، ... وجود دارد. برای مدل کردن ارتباطات دودویی میان زیرمجموعه‌های

یک مجموعه، استفاده شایان توجهی از ابرگراف می‌شود. شکل زیر مثالی جالب از مدل کردن گراف جهتدار

یک مسئله دنیای واقعی به یک ابرگراف را نشان می‌دهد.

ابرگراف در هندسه محاسباتی نیز کاربرد بسیاری دارد. در آن حوزه، هر ابرگراف معادل با range space و

هر ابرپال معادل با یک range فرض شده و آنگاه به حل مسئله اصلی اقدام می‌شود. همچنین در حیطه‌های

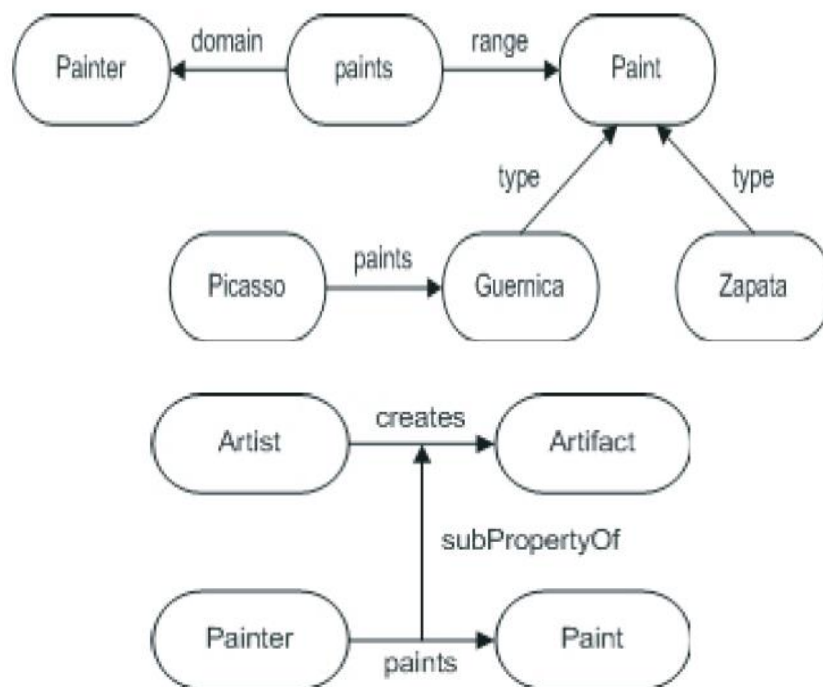
مختلف مدیریتی و نیز مهندسی نرم‌افزار که مطالعه چرخه فعالیتها بسیار مهم است از ابرگراف استفاده

می‌شود. شکل زیر دو گراف RDF ساده را در روند کلی مدلسازی فعالیتها و وقوع و تکرار موجودیتها و نیز

اَعمال مجاز نشان می‌دهد. RDF یک زبان برای چارچوب توصیف منابع^۱ و بیان metadata درباره منابع

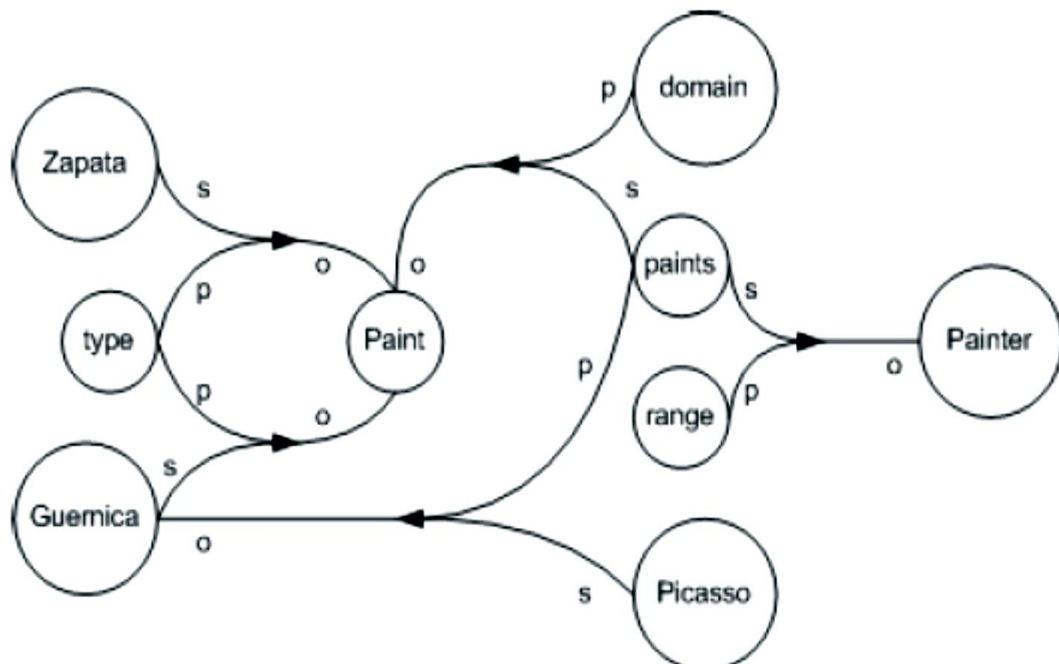
مسئله است.

^۱ Resource Description Framework (RDF)



شکل ۴-۵: توپولوژی توصیفی برای یک مسئله RDF [۳۰]

ابرگراف بدست‌آمده از توصیف روابط میان منابع و رفتار آنها در شکل زیر نشان داده می‌شود:



شکل ۴-۶: مدل کردن مسئله RDF با یک ابر گراف [۳۰]

۴-۳- ارتباط سلسله مراتبی ابر گراف بر مبنای دنبال کننده

آشکارسازی‌های فریم-فریم معین، فرآیند دنبال کردن را بعنوان یک مسئله جستجوی همسایگی‌های متراکم سلسله مراتبی بر روی ارتباط وابستگی مسیرک‌های ابر گراف مطرح می‌کند بگونه‌ای که با توصیف رابطه بین مسیرک‌های تولید شده لایه قبل، بصورت پویا لایه جدید ساخته می‌شود.

نکته قابل ملاحظه اینست که یک آشکارسازی بعنوان یک مسیرک کاهش یافته در نظر گرفته می‌شود که شامل فقط یک پاسخ آشکار شده است. در ابتدا فایل ویدیویی به u قسمت در حوزه زمان تقسیم می‌شود. اجازه دهید Δ_r^l نشان دهنده فاصله زمانی r قسمت در لایه l و T طول کل فریم‌های ویدیو باشد. مجموعه فاصله زمانی u قسمت بصورت زیر نمایش داده می‌شود [۳۷]:

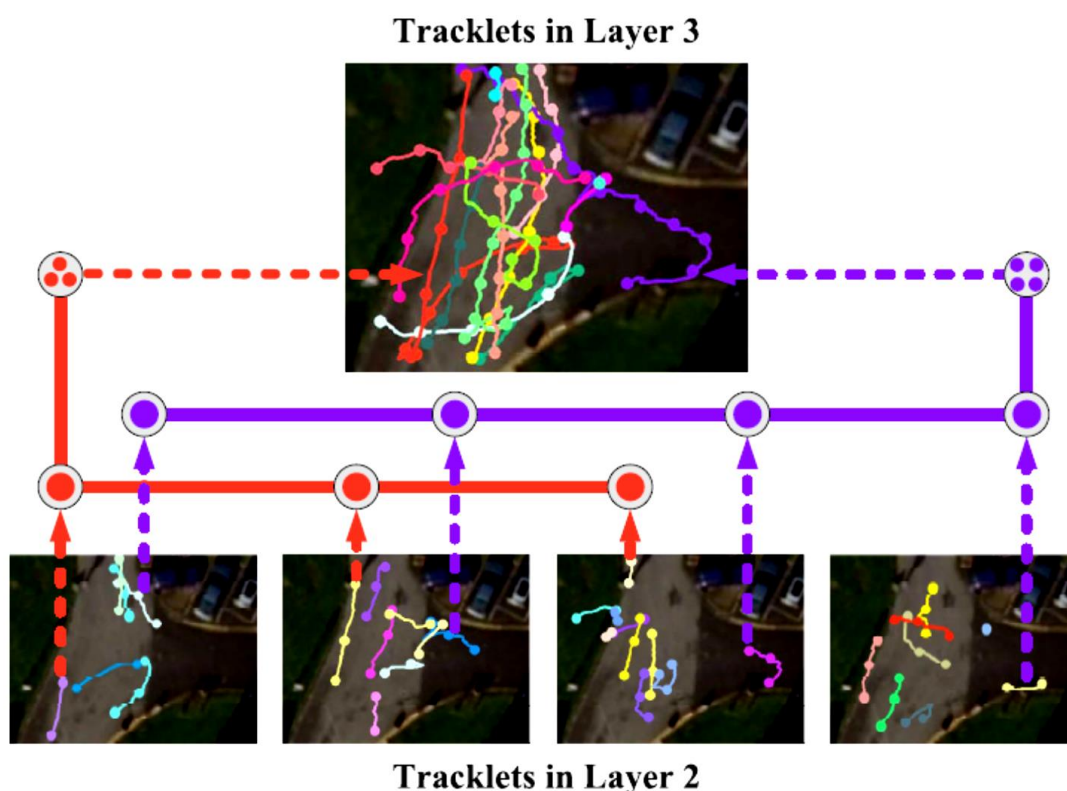
$$\{\Delta_1^l, \Delta_2^l, \dots, \Delta_u^l\} \quad (4-1)$$

Such that $\Delta_i^l \cap \Delta_j^l = \emptyset$ and $\bigcup_{i=1}^u \Delta_i^l = T$

گراف نزدیکی روابط در هر قسمت در همان لایه ساخته می‌شود. اگر $G_r^l = (V_r^l, E_r^l)$ نشان دهنده گراف ساخته شده در قسمت r ام و لایه l ام باشد آنگاه V_r^l و E_r^l به ترتیب مجموعه راس‌ها و یال‌های گراف هستند. مجموعه راس‌های گراف برابر خواهند شد با $V_r^l = \{v_{i,r}^l\}_{i=1}^n$ که i امین راس و n تعداد راس‌ها در گراف است. هر راس در G_r^l متناظر با یک مسیرک در همان قسمت است. مجموعه متناظر مسیرک‌ها بصورت $\mathcal{T}_r^l = \{\mathcal{T}_{i,r}^l\}_{i=1}^n$ تعریف می‌شود. یال‌های گراف رابطه بین مسیرک‌های مختلف را توصیف می‌کند و وزن‌های آنها احتمال متعلق بودن مسیرک‌ها به یک هدف را نشان می‌دهد. در ادامه $v_{i,r}^l$ و $\mathcal{T}_{i,r}^l$ را چنانکه بتوان بجای یکدیگر بکار برد برای نشان دادن i امین قسمت در لایه l استفاده می‌کنیم. متفاوت از کارهای قبل که فقط رابطه‌های دوبدو را مورد بررسی قرار می‌دادند، در این روش روابط مرتبه بالا در میان چندین مسیرک در گراف ارتباط برای تعدادی از لایه‌های ابتدایی یکپارچه می‌شود بدین معنی که یال‌ها در گراف بیشتر از دو راس را بهم وصل می‌کنند. در این روش حرکت و همواری مسیر برای اطمینان از کارایی روش

دنبال کردن بکارگرفته شده است. الگوریتم جستجوی همسایگی‌های متراکم نیز برای مسئله آشکارسازی همسایگی‌های متراکم در هر دو گراف و ابر گراف استفاده شده است تا مسیرک بلندتری را بوسیله چسباندن مسیرک‌ها در هر همسایگی متراکم آشکار شده تولید کند. قابل توجه اینکه تمامی گراف در همه قسمت‌ها به یک صورت پردازش می‌شود.

بعد از کامل شدن مسئله جستجوی همسایگی‌های متراکم در همه قسمت‌های یک لایه ، لایه‌های نزدیک بهم باهم ترکیب می‌شوند تا تقسیم‌بندی جدیدی در لایه بعد شکل بگیرد. سپس ساختار گراف برای هر قسمت از لایه بعد ساخته می‌شود و آشکارسازی همسایگی‌های متراکم روی این گراف برای تولید مسیرک بلندتر صورت می‌پذیرد. این دو مرحله تکرار می‌شوند تا زمانی که فقط یک قسمت در لایه باقی بماند. در نهایت جستجوی همسایگی متراکم بر روی گراف وابستگی ساخته شده در آن قسمت اجرا می‌شود تا مسیر حرکت نهایی اهداف بدست آید. بعنوان مثال فرآیند بهینه سازی از مرحله ۲ به ۳ در شکل ۴-۷ نشان داده شده است.



شکل ۴-۷: بازیابی مسیرکها در یک فرآیند بهینه سازی از لایه سوم به لایه دوم که ۴ قسمت تصویری باهم ترکیب شده‌اند و مسیرکهای نهایی در لایه سوم ساخته شده است. خطهای آبی و قرمز ترکیب بهینه مسیرکها را نشان می‌دهد [۳۷].

۴-۴- ساختار گراف وابستگی غیر جهت دار

برای یک قسمت جدا شده از فایل رشته ویدیویی، ما ارتباط کلی گراف وابستگی را با توجه به $G = (V, E)$ ایجاد کردیم که یک گراف کامل است و روابط بین مسیرکهای مختلف را توصیف می‌کند. اگر $V = \{v_1, v_2, \dots, v_n\}$ مجموعه نودهای گراف باشد و E مجموعه یالهای گراف باشد آنگاه $E \subset V \times V \times \dots \times V$ که m بار مجموعه V ضرب می‌شود و m تعداد راس‌های درگیر در هر یال است. بطور واضح گراف ساخته شده یک ابر گراف است اگر $m > 2$ باشد و به سمت یک گراف کلی میل می‌کند اگر $m = 2$ شود. در اینجا برای نشان دادن راس های m تایی درگیر در یالهای گراف از $e^m = (v_1, v_2, \dots, v_n)$ استفاده می‌کنیم. آرایه شباهت $\mathcal{A}$ گراف G یک آرایه کاملاً متقارن است بگونه‌ای که عناصر آن نشان دهنده احتمال تعلق مسیرک ها در e^m به یک هدف یکسان است. مقادیر شباهت در آرایه بصورت زیر تعریف می‌شود:

$$\begin{cases} \mathcal{A}(e^m) = 0 & \text{if } e^m \notin E \\ \mathcal{A}(e^m) \geq 0 & \text{otherwise} \end{cases} \quad (2-4)$$

طراحی نحوه محاسبه مقادیر شباهت در گراف غیر جهت دار یک نقش اساسی را در الگوریتم بازی می‌کند بطوریکه میزان احتمال متعلق بودن مسیرک‌ها را به یک هدف یکسان نشان می‌دهد. مقادیر شباهت یال‌ها $\mathcal{A}(e^m)$ بوسیله ۳ فاکتور محاسبه می‌شود: شباهت، حرکت و همواری مسیر. فاکتور شباهت ظاهری نشان دهنده همسانی ظاهری در بین مسیرک‌هاست، فاکتور حرکت نشان دهنده پیوستگی حرکت بین مسیرک‌ها می‌باشد و فاکتور همواری مسیر نشان دهنده همواری فیزیکی مسیر حرکت در مسیرک‌های یکپارچه شده می‌باشد. بنابراین مقادیر شباهت یال‌ها (مقادیر وزن) بصورت زیر محاسبه می‌شود [۳۷]:

$$\mathcal{A}(e^m) = w_1 \mathcal{A}_a(e^m) + w_2 \mathcal{A}_m(e^m) + w_3 \mathcal{A}_s(e^m) \quad (3-4)$$

که $\mathcal{A}_a(\cdot)$ ، $\mathcal{A}_m(\cdot)$ و $\mathcal{A}_s(\cdot)$ به ترتیب شباهت ظاهر، حرکت و همواری مسیر هستند. w_1 ، w_2 و w_3 نیز پارامترهای بالانس این سه فاکتور می‌باشند. واضح است که اگر در حوزه زمان دو مسیرک باهم همپوشانی داشته باشند متعلق به یک هدف نیستند بنابراین $\mathcal{A}(e^m) = 0$. بنابراین ما فقط مواردی را محاسبه می‌کنیم که هیچکدام از دو مسیرک با هم هم پوشانی نداشته باشند.

۴-۴-۱- شباهت (وابستگی) ظاهری

ابهام زدایی از ظاهر یک شی و تمیز دادن از پس زمینه و دیگر اشیا یکی از بخش‌های بسیار سخت و مشکل می‌باشد. در روش ارائه شده دو نوع ویژگی در فریم اول و فریم آخر مسیرک برای توصیف ظاهر آن شی مورد استفاده قرار گرفته است. یعنی ویژگی هیستوگرام رنگ و ویژگی هیستوگرام گرادیان شکل. در هیستوگرام رنگ RGB برای هر کانال ۸ خانه استفاده شده است و در هیستوگرام گرادیان شکل ۳۶ اندازه در نظر گرفته شده است. بدون از دست دادن کلیت مسئله، فرض می‌شود مسیرک v_i قبل از مسیرک v_j در حوزه زمان ظاهر شود آنگاه شباهت ظاهری بصورت زیر محاسبه می‌شود [۳۷]:

$$\mathcal{A}_a(e^m) = \sum_{v_i, v_j \in e^m} \sum_{k=1,2} \lambda_k H(\tilde{f}_k(v_i), \hat{f}_k(v_j)) \quad (4-4)$$

که $\tilde{f}_1(v_i)$ و $\tilde{f}_2(v_i)$ هیستوگرام رنگ و شکل آخرین فریم و $\hat{f}_1(v_i)$ و $\hat{f}_2(v_i)$ هیستوگرام شکل و رنگ اولین فریم از مسیرک مربوطه می‌باشند. H نمایانگر فاصله کسینوسی بین دو بردار ویژگی و λ_1, λ_2 پارامترهای بالانس را نشان می‌دهند.

۴-۴-۲- شباهت حرکتی

یکی از تعاریف درست در فرآیند دنبال کردن آنست که حرکت هدف بصورت آهسته و متناسب با میزان سرعت فریم‌ها در زمان باشد. همچنین نکته قابل توجه دیگر اینست که می‌توان سرعت یک هدف را در یک بازه زمانی کوچک ثابت فرض کرد. بنابراین شباهت حرکتی در واقع پیوستگی حرکت مسیرک‌ها را منعکس می‌کند که بصورت زیر تعریف می‌شود [۳۷]:

$$\mathcal{A}_m(e^m) = \sum_{v_i, v_j \in e^m} S_m(v_i, v_j) \quad (5-4)$$

که S میزان شباهت حرکتی بین مسیرک v_i و v_j است. در لایه اول بدلیل اینکه فقط یک پاسخ آشکارسازی وجود دارد بنابراین در هرکدام از مسیرک‌ها اطلاعات شامل شباهت حرکتی وجود ندارد بنابراین مقدار شباهت حرکتی در لایه اول برابر صفر قرار می‌گیرد.

ما فرض می‌کنیم مسیرک v_i قبل از v_j ظاهر شود. اگر $D_{v_i} = \{d_s^{v_i}\}_{s=1}^{|T_{v_i}|}$ مجموعه آشکارسازی شده از مسیرک v_i باشد، d را s امین آشکارسازی در مسیرک و T را تعداد آشکارسازی‌ها در v_i در نظر می‌گیریم. تابع حرکت خطی را بصورت زیر تعریف می‌کنیم [۳۷]:

$$\mathcal{P}(l(.), \Delta t, \vec{v}) = l(.) + \Delta t. \vec{v} \quad (6-4)$$

که $l(.)$ مکان آشکارسازی شده است، Δt فاصله زمانی و $\vec{v}$ سرعت ثابت در حوزه زمان است. اگر $t(d_s^{v_i} - d_s^{v_j})$ فاصله زمانی بین آشکارسازی‌های $d_s^{v_i}$ و $d_s^{v_j}$ باشد آنگاه با توجه به این حقیقت که اندازه کرولیشن بین فریم‌های نزدیک زیاد و بین فریم‌های دور کم می‌باشد، ما فقط از $\mathcal{T}$ آشکارسازی در تعدادی از فریم‌های آخر v_i و تعدادی از فریم‌های اول v_j برای محاسبه میزان شباهت حرکتی بین آنها استفاده می‌کنیم. برای کاهش تاثیر نویز انحراف‌های هر دو پیش بینی رو به عقب v_i و پیش بینی رو به جلوی v_j برای اندازه‌گیری پیوستگی حرکت بین دو مسیرک مورد استفاده قرار می‌گیرد [۳۷].

اگر

$$l_{p,q}^{v_i}(d_{l_i}^{v_i}) = \mathcal{P}(l(d_{l_i}^{v_i}), t(d_{l_i}^{v_i} - d_{l_i}^{v_j}), \frac{l(d_{l_i}^{v_i}) - l(d_{l_i}^{v_j})}{t(d_{l_i}^{v_i} - d_{l_i}^{v_j})}) \quad (۷-۴)$$

موقعیت شروع پیش‌بینی شده از آشکارسازی $d_{l_i}^{v_i}$ با سرعت متوسط بین موقعیت‌های $d_{l_i}^{v_i}$ و $d_{l_i}^{v_j}$ باشد، $l_i = |\mathcal{T}_{v_i}|$ برابر تعداد پاسخ‌های آشکارسازی شده در مسیرک $\mathcal{T}_{v_i}$ باشد آنگاه شباهت حرکتی بین دو مسیرک v_i و v_j بصورت زیر محاسبه می‌شود:

$$S_m(v_i, v_j) = \sum_{p=1}^{\mathcal{T}} \sum_{q=1}^{p-1} \exp\left(-\frac{\|l(d_{l_i}^{v_i}) - l_{p,q}^{v_j}(d_{l_i}^{v_j})\|^2}{\sigma_m^2}\right) + \sum_{p=l_i-\mathcal{T}}^{l_i-1} \sum_{q=p+1}^{l_i} \exp\left(-\frac{\|l(d_{l_i}^{v_j}) - l_{p,q}^{v_i}(d_{l_i}^{v_i})\|^2}{\sigma_m^2}\right) \quad (۸)$$

۴-۴-۳- شباهت همواری مسیر

مسیرهای حرکت هدف در حوزه زمان-مکان بصورت پیوسته و هموار است و از این روی اطلاعات موثر و مفیدی برای اندازه‌گیری میزان اطمینان از متعلق بودن مسیرک‌های در e^m به یک هدف یکسان را مهیا می‌کند. در ابتدا مسیرک‌ها در ابر یال e^m با توجه به زمان ظاهر شدنشان باهم ترکیب می‌شوند تا مسیرک

فرضی T_e^m بدست آید. پاسخ‌های آشکارسازی مرتب شده در مسیر T_e^m با توجه به افزایش زمان برابر زیر در نظر گرفته می‌شود [۳۷]:

$$U_{i=1}^m \{d_j^{v_i}\}_j^{l_i} = 1 \quad (9-4)$$

که $l_i = |T_{v_i}|$ نشان دهنده تعداد پاسخ‌های آشکار شده در مسیر است. ما بعضی آشکارسازی‌ها در مسیر را با گام برابر دلتا برای معین کردن مجموعه نقاط آشکارسازی شده مناسب بصورت زیر انتخاب می‌کنیم:

$$D_{em}^f = \{d_i\}_{i=1+k\delta, k \in \mathbb{N}}^{1 \leq i \leq \sum |T_{v_i}|} \quad (10-4)$$

مجموعه نقاط آشکارسازی باقی‌مانده برابر است با:

$$D_{em}^r = D_{em} \setminus D_{em}^f \quad (11-4)$$

ما از D_{em}^f برای تعیین مسیر حرکت مناسب $\hat{T}_e^m$ از مسیر یکپارچه شده فرضی با بکارگیری الگوریتم درونیایی اسپلاین درجه سوم استفاده می‌کنیم و انحراف بین نقاط آشکار شده در مجموعه D_{em}^r و نقاط مسیر حرکت $\hat{T}_e^m$ با اندیس زمانی مشابه را محاسبه می‌کنیم. بنابراین *smooth affinity* ابر یال em بصورت زیر محاسبه می‌شود:

$$\mathcal{A}_s(e^m) = \exp \left(- \sum_{d_i \in D_{em}^r} \frac{\|l(d_i) - l(\hat{T}_e^m(t(d_i)))\|^2}{\sigma_s^2} \right) \quad (12-4)$$

که $l(d_i)$ موقعیت پاسخ‌های آشکار شده d_i است، $t(d_i)$ زمان ظاهر شدن پاسخ‌های آشکار شده d_i است و $\hat{T}_e^m(t(d_i))$ نشان دهنده مسیر حرکت $\hat{T}_e^m$ در زمان $t(d_i)$ است.

۴-۵- جستجوی همسایگی‌های متراکم مسیرک‌ها

بعد از ساختن گراف ارتباط مسیرک‌ها، ما همسایگی‌های متراکم را بر روی آن آشکار می‌کنیم. مسئله اصلی در آشکارسازی همسایگی‌های متراکم چگونگی تعیین تعداد همسایگی‌ها و تعداد نودها در هر همسایگی است. در اینجا تعداد همسایگی‌ها و تعداد نودها در هر همسایگی بعنوان یک متغیر پنهان در فرآیند بهینه

سازی در نظر گرفته شده است که بوسیله ماکزیمم کردن متوسط مقادیر شباهت همسایگی‌ها بدست می‌آید. در این روش چندین همسایگی متراکم بصورت کاملاً موفق کشف می‌شود. سپس مسیرک نهایی در هر قسمت می‌تواند بوسیله متصل کردن مسیرک‌ها در هر همسایگی بطور صحیح بدست آید. جهت اطمینان از آشکار شدن تمام همسایگی‌های متراکم، هر راس در گراف بعنوان یک نقطه شروع و جستجوی همسایگی-های متراکم خودش در نظر گرفته شده است. اگر مسیرک متعلق به یک همسایگی متراکم واقعی باشد، وابستگی بین آن و دیگر نودهای قرار گرفته در نزدیکی آن معمولاً زیاد است. بنابراین همسایگی‌های متراکم آن پیدا خواهند شد. در مقابل اگر مسیرک رابطه ضعیفی با دیگر مسیرک‌ها داشته باشد، میزان وابستگی بین آنها کم خواهد بود که نشان می‌دهد آن متعلق به هیچکدام از همسایگی‌های متراکم منسجم نیست. از این رو مسیرک بعنوان یک مثبت کاذب تلقی می‌شود و آن را در طبقه‌بندی نهایی نادیده می‌گیریم. برای نقطه شروع vp ، هدف پیدا کردن همسایگی‌های متراکم آن یعنی $\mathcal{N}(v_p)$ است به گونه‌ای که شامل ماکزیمم متوسط مقادیر وابستگی شود. مسئله بهینه سازی بصورت زیر فرمول بندی می‌شود [۳۷]:

$$\mathcal{N}^*(v_p) = \arg \max_{\mathcal{N}(v_p)} \mathcal{C}(v_p \cup \mathcal{N}(v_p)) \quad (۱۳-۴)$$

$$s. t. \mathcal{N}(v_p) \subset V, v_p \notin \mathcal{N}(v_p), |\mathcal{N}(v_p)| = k$$

که $\mathcal{C}(v_p \cup \mathcal{N}(v_p))$ تابع اندازه‌گیری وابستگی (شباهت) است که توزیع وابستگی در گراف را نشان می‌دهد. فرض کنید $U = \{v_p \cup \mathcal{N}(v_p)\}$ نشان دهنده مجموعه شامل راس vp و k راس همسایگی آن باشد. بنابراین زیرمجموعه $U \subset V$ شامل $k+1$ راس خواهد بود. اگر $\mathcal{Y} \in \mathbb{R}^n$ بردار شاخص زیرمجموعه U باشد یعنی:

$$\begin{cases} y_i = 1 & \text{if } v_i \in U \\ y_i = 0 & \text{otherwise} \end{cases} \quad (۱۴-۴)$$

بنابراین موضوع فرمول ۱۳-۴ بصورت $\sum_{i=1}^n y_i = k + 1, y_i \in \{0, 1\}, y_p = 1$ تغییر می‌کند.

دو شرط اول بیان می‌کنند که وجود دارد $k+1$ مسیرک متعلق به یک هدف بطوریکه بوسیله حل $\mathcal{Y}$ نشان داده شده است و شرط آخر بیان می‌کند که راه‌حل باید شامل مسیرک vp باشد. فرض می‌کنیم Eu مجموعه

یال‌های متناظر با مجموعه رئوس U باشد. اگر مسیرک‌ها در U متعلق به یک هدف یکسان باشند، اغلب یال-ها در E_U مقادیر وابستگی بزرگی خواهند داشت. البته مقادیر وابستگی کلی مجموعه یال E_U بصورت زیر محاسبه می‌شود:

$$\mathcal{C}(U) = \sum_{e^m \in E_U} \mathcal{A}(e^m) \quad (15-4)$$

در فرآیند دنبال کردن مقادیر وابستگی در ابر گراف غیر منفی می‌باشند، یعنی $\mathcal{A}(e^m) \geq 0$. واضح است که $\mathcal{C}(U)$ بطور معمول با تعداد راس‌ها در زیرمجموعه U افزایش می‌یابد. از این روی آن منطقی‌تر است تا متوسط مقادیر وابستگی را برای توصیف میزان اطمینان از همسایگی‌های متراکم نسبت به مقادیر کلی وابستگی استفاده کنیم. بطوریکه بتوان انحراف اندازه بین همسایگی‌های متراکم را کنترل کرد. از آنجایی که $\sum_{i=1}^n y_i = k+1$ ، $(k+1)^m$ جمع‌وند در $\tilde{\mathcal{C}}(U)$ وجود دارد. متوسط مقادیر $\mathcal{C}(U)$ انتخاب می‌شود تا بتوانیم همسایگی‌های متراکم درست را نشان دهیم بدین صورت:

$$\mathcal{C}(U) = \frac{1}{(k+1)^m} \tilde{\mathcal{C}}(U) = \sum_{e^m \in E_U} \mathcal{A}(e^m) \frac{y_1}{k+1} \dots \frac{y_m}{k+1} \quad (16-4)$$

سپس مسئله بهینه‌سازی در فرمول ۱۳ می‌تواند مانند زیر بیشتر ساده شود:

$$\begin{aligned} \max_x g(x) &= \sum_{e^m \in E_U} \mathcal{A}(e^m) x_1 \dots x_m \\ \text{s.t. } \sum_{i=1}^n x_i &= 1, x_i \in \{0, \epsilon\}, x_p = \epsilon. \end{aligned} \quad (17-4)$$

که $x_i = y_i / (k+1)$ و $\epsilon = \frac{1}{k+1}$. در واقع یک مسئله بهینه‌سازی ترکیبی است. بنابراین برای ساده‌سازی بیشتر $0 < x_i < \epsilon$ در نظر می‌گیریم.

۴-۶- روش پیشنهادی جهت بهبود الگوریتم:

۴-۶-۱- کنترل تصادم با اضافه کردن راس فرضی :

در برخی از فریم‌ها ممکن است آشکارسازی بطور دقیق برای یک شخص بدلیل مخفی شدن قسمتی از بدن یا کل بدن در پشت مانع‌های موجود در تصویر رخ نداده باشد. لذا در این هنگام با گم شدن لحظه‌ای شخص در قسمت تصویری مربوطه مواجهه خواهیم شد. برای حل کردن این مشکل پیشنهاد می‌کنیم یک راس فرضی در آخرین موقعیت وجود شخص مورد نظر به خوشه اضافه شود. اگر Vs راس فرضی درست در مقابله با داده‌های نادرست حاصل از آشکارسازی باشد، می‌توان با توجه به مدت زمان کوتاه تصادم در هنگام راه رفتن عابر سرعت آن را ثابت در نظر گرفت. لذا با محاسبه سرعت عابر در چند فریم منتهی به پنهان شدن، می‌توان خط سیر آن طی چند فریمی که عابر دچار تصادم یا مشکل مخفی شدن شده است، ترسیم کرد و این خط سیر را به اولین فریم دارای پاسخ درست در آشکارسازی متصل نمود.



شکل ۴-۸: تصحیح خطای خط سیر اهداف در هنگام مخفی شدن عابر در پشت مانع با اضافه کردن راس فرضی

۴-۷- نتایج پیاده‌سازی

الگوریتم پیشنهادی بر روی پایگاه داده ParkingLot با خصوصیات تراکم عابر پیاده بالا در تصاویر و برخوردهای زیاد و با نرخ ۶ فریم بر ثانیه تست شده است. پایگاه داده ParkingLot شامل ۱۰۰۰ فریم از صحنه‌های شلوغ و پر تردد می‌باشد که در آن عابران پیاده بصورت موازی در یک محوطه پارکینگ در حال

راه رفتن می‌باشند. از جمله چالش‌های موجود در این پایگاه می‌توان به تصادم و برخورد زیاد، آشکارسازی-های همراه با خطا و حرکت موازی عابرانی با ظاهر شبیه بهم اشاره کرد. پارامترهای مرجع استفاده شده در این تحقیق در زیر بیان شده اند. پارامترهای وزن در مقادیر شباهت بصورت $w_1=0.6$, $w_2=0.2$, $w_3=0.2$ در نظر گرفته شده‌اند. پارامترهای توازن بین هیستوگرام رنگ RGB و هیستوگرام گرادیان شکل $\lambda_1 = 0.1$ و $\lambda_2 = 0.04$ فرض شده‌اند. مقادیر شباهت حرکتی نیز برابر $\sigma_m^2 = 5$ و $\sigma_s^2 = 4$ قرار داده شده اند. هر ۶-۸ فریم ترکیب شده‌اند تا بخش‌های لایه اول تولید شود و ۳-۵ بخش برای تولید لایه‌های بالاتر مورد استفاده قرار گرفته است.



قسمت اول گروه تصاویر



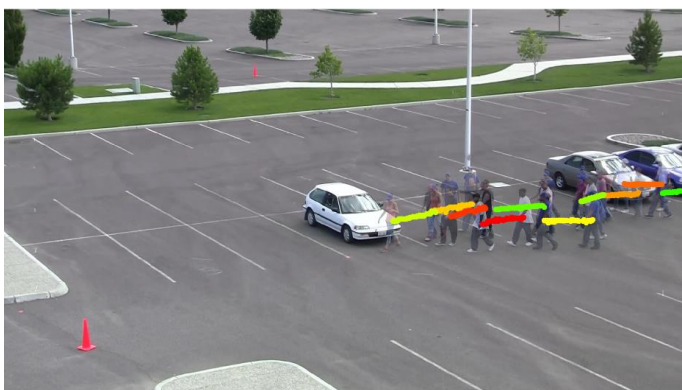
قسمت دوم گروه تصاویر



قسمت سوم گروه تصاویر



قسمت چهارم گروه تصاویر



قسمت پنجم گروه تصاویر



قسمت ششم گروه تصاویر

شکل ۴-۹ : نتایج آنالیز بخش اول پایگاه داده ParkingLot

علاوه بر نمایش تصویری نتایج، دو معیار عددی نیز برای تحلیل روش پیشنهادی بکار گرفته شده است. معیار نرخ فراخوانی درست برای سنجش توانایی روش پیشنهادی برای ردیابی اهداف در هنگام تصادم (Recall یا Rate) و معیار نرخ تشخیص غلط (False Alarm) یا FAR برای بررسی تأثیر تصادم اهداف با یکدیگر در نظر گرفته شده است این دو معیار بصورت معادلات زیر تعریف می گردند:

$$ReR = \frac{\text{تعداد تخمین درست}}{\text{تعداد اهداف واقعی در کل فریم ها}}$$

$$FAR = \frac{\text{تعداد تخمین غلط}}{\text{تعداد کل تخمین ها}}$$

جدول ۴-۱ نتایج مربوط به مجموعه داده‌گان یک و دو را بر اساس این دو معیار نشان می‌دهد

شماره گروه تصاویر	شماره فریم	مقدار ReR	مقدار FAR
گروه تصاویر ۱	۵۰-۱	۱	۰
گروه تصاویر ۲	۱۰۰-۵۱	۰٫۸۳	۰٫۱۶
گروه تصاویر ۳	۱۵۰-۱۰۱	۱	۰
گروه تصاویر ۴	۲۰۰-۱۵۱	۱	۰
گروه تصاویر ۵	۲۵۰-۲۰۱	۱	۰
گروه تصاویر ۶	۳۰۰-۲۵۱	۰٫۹۲۸	۰٫۰۷۱

جدول ۴-۱: نتایج حاصل از آنالیز ۶ قسمت اول پایگاه داده ParkingLot

نتایج آزمایشات انجام شده بر روی داده‌های پایگاه داده ذکر شده نشان می‌دهند که روش ارائه شده است، صرف نظر از میزان پیچیدگی حرکت اهداف و در شرایط وجود تصادم خط سیر مربوط به اهداف را بخوبی مشخص کرده است.

فصل پنجم

نتیجه‌گیری و پیشنهادات

۵-۱- جمع‌بندی و نتیجه‌گیری

هدف رهگیری همزمان چند شی، تخمین وضعیت چندین هدف با حفظ هویت آن‌ها و با در نظر گرفتن تغییرات حرکت و ظاهر آن‌ها با گذشت زمان می‌باشد. در یک محیط مختلط و شلوغ، این موضوع به خصوص بواسطه تصادم پی در پی ناشی از بهم ریختگی، ظاهر شبیه به یکدیگر سایر اهداف و ... بیشتر چالش برانگیز می‌شود. به هر حال، به هنگام وقوع تصادم‌های بلند مدت و پیوسته، تشخیص و جداسازی اهداف مختلف دشوار بوده و همین امر محدودیت‌هایی زیادی را در پیاده‌سازی روش‌های معمول ایجاد می‌کند. بطوریکه در هنگام آنالیز این تصویرها توسط روش‌هایی نظیر فیلتر کالمن و یا فیلتر ذره‌ای به هیچ عنوان جواب قابل قبولی حاصل نمی‌شود.

در سال‌های اخیر به دلیل توسعه روش‌های مبتنی بر آشکارسازی و تشخیص مطمئن اشیا حتی در محیط‌ها و صحنه‌های شلوغ، عملکرد روش‌های رهگیری توسط آشکارسازی بهبود قابل توجهی داشته است. به طور کلی، روش‌های رهگیری توسط آشکارسازی، با استفاده از اطلاعات جمع‌آوری شده از آشکارسازها مکان هندسی^۱ مربوطه را ایجاد می‌کند.

این روش‌ها معمولاً از تشخیص کلی قاب‌ها به منظور به هم متصل کردن مکان‌های هندسی کوچک^۲ جدا از هم^۳ استفاده می‌کنند. با در دست داشتن آشکارسازی‌های صورت گرفته برای یک مجموعه از فریم‌ها، مکان‌های هندسی کوچکتر با به هم پیوستن آشکارسازی‌ها پدید آمده و سپس مکان‌های هندسی کوچکتر به منظور ساخت مکان هندسی طولانی‌تر به یکدیگر پیوند داده می‌شوند. بنابراین در این دیدگاه، تجمیع کلی^۴ بسیار اثرگذار و حائز اهمیت است.

^۱. Trajectory

^۲. Tracklet

^۳. Fragmented

^۴. Global association

روش پیشنهادی با فرض آشکارسازی فریم به فریم مساله رهگیری چند هدف همزمان را با جستجوی همسایگی‌های متراکم در ابر گراف مدل می‌کند.

برای اطمینان از رهگیری صحیح اهداف سه معیار شباهت ظاهری، شباهت حرکتی و شباهت همواری مسیر در نظر گرفته شده است.

معیار کلی شباهت از ترکیب این سه معیار با ضرایب وزنی متفاوت بدست می‌آید. این ضریب نمایانگر وزن ابريال‌ها است. راس‌ها نیز بعنوان مسیرک‌ها در نظر گرفته می‌شوند و ارتباط آنها با ابريال‌های وزندار برقرار می‌شود.

جستجوی همسایگی‌های متراکم به جای روش‌های معمول جستجوی نزدیکترین همسایگی مد نظر قرار گرفته است. زیرا در الگوریتم‌های معمول فقط ارتباط نقطه مورد نظر با نزدیکترین همسایگی‌ها مورد بررسی قرار می‌گیرد حال آنکه ممکن است وجود داده‌های پرت عناصری کاملاً نامرتبط را به یکدیگر مرتبط نشان دهد.

بطور کلی پذیرفته شده که اهداف متعلق به یک خوشه شرایط بهم پیوستگی داخلی مشخصی را برآورده می‌کنند در حالیکه اگر متعلق به یک خوشه نباشند معمولاً این شرط برآورده نمی‌شود. اغلب روش‌های خوشه‌بندی موجود مبتنی بر جزءبندی می‌باشند از قبیل k -NN، k -means^۲، کلاسه بندی فضایی و انتشار وابستگی^۶. این روش‌ها در مفهوم از یک فرض بهره می‌برند: هر نقطه داده می‌بایست متعلق به یک خوشه باشد. این فرض مسئله را تا حدود بسیار زیادی ساده می‌کند از این روی که ما نیاز نداریم تا در مورد یک نقطه داده قضاوت کنیم که آیا داده پرت است یا خیر، که بسیار چالش برانگیز است. به هر حال این فرض عملکرد بدی را در این روش‌ها سبب می‌شود هنگامی که تعداد زیادی داده پرت وجود داشته باشد همانگونه که در بسیاری از کاربردهای در دنیای واقعی با آنها سروکار داریم.

از این روی در الگوریتم پیشنهادی معیار قضاوت اینکه چندین هدف متعلق به یک خوشه هستند یا خیر بوسیله ارتباطی فراتر از ارتباط دودو بیان می‌شود و ارتباط چندتایی داده‌ها با یکدیگر مد نظر قرار می‌گیرد که بوسیله وزن یال‌های یک ابرگراف وابستگی مدل می‌شوند. این الگوریتم علاوه بر مقایسه و سنجش معیار

مشابهت بین نقطه داده مورد نظر و همسایگی‌های آن، ارتباط تمامی همسایگان آن نقطه را نیز بصورت دوبدو مورد تجزیه و تحلیل قرار می‌دهد و از این روی قابلیت جداسازی اهدافی با ظاهری شبیه بهم را در هنگام تصادم فراهم می‌کند.

علاوه بر این، در روش پیشنهادی در هر تکرار، مکان‌های هندسی به دست آمده به یکدیگر پیوند داده می‌شوند که با استفاده از ضریب همواری مسیرک‌ها^۱ می‌توان مدیریت مکان اهداف در مواردی که با تصادم اهداف و یا آشکارسازی غیر قابل اعتماد مواجه می‌شویم از آن بعنوان تصحیح کننده بهره ببریم.

همانطور که پیشتر توضیح داده شد، مراحل اصلی روش پیشنهادی تجمیع‌های محلی و سراسری می‌باشد. در هر دو گام، مدلسازی ظاهر اهداف به منظور تجمیع مسیرک‌ها و آشکارسازی‌های صورت گرفته از اهداف یکسان و در عین حال تشخیص اهداف شبیه به هم، اثرگذار و حائز اهمیت می‌باشد. برای نیل به این هدف، از ساختار تحلیل سلسله مراتبی استفاده می‌کنیم بدین صورت که در هر لایه ابر گراف یا گراف مخصوص به همان لایه ساخته شده و آنالیز جستجوی همسایگی‌های متراکم برای همان لایه انجام خواهد شد.

این در حالی است که بیشتر روش‌های ارائه شده قبلی در بحث رهگیری، تنها آنالیزهای بصورت محلی بوده و ساختار ابرگراف/گراف بصورت پویا تغییر نمی‌کرده است.

به هر حال روش‌های رهگیری مبتنی بر تجمیع محلی^۲ صرف، قادر به تولید مکان‌های هندسی جدا از هم می‌باشند و احتمال وقوع خطا در آن‌ها در اثر وقوع تصادم بسیار زیاد می‌باشد. این موضوع ناشی از آن است که روش‌های یاد شده تنها از اطلاعات دو فریم متوالی استفاده می‌کنند.

این در حالیست که روش پیشنهادی علاوه بر در نظر گرفتن ارتباطات محلی فریم‌ها آنالیز آنها را در لایه-های بالاتر و بصورت سراسری نیز انجام می‌دهد که در حل مشکل تصادم در مکان‌های شلوغ بسیار موثر می‌باشد.

^۱. Tracklets

^۲. local association-based-tracking methods

۵-۲- پیشنهادات برای ادامه تحقیق

(۱) تحلیل ابرگراف با استفاده از پیدا کردن کوچکترین ماتریس کلیک در یک قسمت تصویری به

منظور افزایش دقت

(۲) ترکیب فرآیند آشکارسازی با فرآیند رهگیری به منظور ساده‌سازی و کاهش زمان آنالیز جهت

بکارگیری آنلاین سیستم

(۳) بکارگیری روش Deep Learning جهت افزایش دقت در خوشه‌بندی اهداف

(۴) بکارگیری هیستوگرام رنگ در فضاهای رنگ دیگر از جمله $L*a*b$ برای افزایش دقت در تمایز

اهداف مشابه

(۵) بکارگیری معیار بافت برای و ترکیب آن با سایر معیارها برای افزایش قابلیت رهگیری در هنگام

حرکت اهداف با فاصله بسیار نزدیک بهم

- [١] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *CVPR*, pp ٨٨٦-٨٩٣, ٢٠٠٥.
- [٢] B. Wu and R. Nevatia. Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors. *IJCV*, ٧٥(٢):٢٤٧-٢٦٦, ٢٠٠٧.
- [٣] A. Andriyenko, S. Roth, and K. Schindler. Continuous Energy Minimization for Multi-Target Tracking. *IEEE TPAMI*, vol. ٣٥, no. ١, ٢٠١٤.
- [٤] Alan J Lipton, Hironobu Fujiyoshi, and Raju S Patil. Moving target classification and tracking from real-time video. In *Applications of Computer Vision*, ١٩٩٨. WACV'٩٨. Proceedings., Fourth IEEE Workshop on, pp ٨-١٤. IEEE, ١٩٩٨.
- [٥] Chris Stauffer and W Eric L Grimson. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition*, ١٩٩٩. IEEE Computer Society Conference on., volume ٢. IEEE, ١٩٩٩.
- [٦] Ya Liu, Haizhou Ai, and Guang-you Xu. Moving object detection and tracking based on background subtraction. In *Multispectral Image Processing and Pattern Recognition*, pages ٦٢-٦٦. International Society for Optics and Photonics, ٢٠٠١.
- [٧] Changick Kim and Jenq-Neng Hwang. Fast and automatic video object segmentation and tracking for content-based applications. *Circuits and Systems for Video Technology*, IEEE Transactions on, ١٢(٢):١٢٢-١٢٩, ٢٠٠٢.
- [٨] Shahbe Mat Desa and Qussay A Salih. Image subtraction for real time moving object extraction. In *Computer Graphics, Imaging and Visualization*, ٢٠٠٤. CGIV ٢٠٠٤. roceedings. International Conference on, pages ٤١-٤٥. IEEE, ٢٠٠٤.
- [٩] Robert T Collins, Alan Lipton, Takeo Kanade, Hironobu Fujiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt, et al. A system for video surveillance and monitoring, volume ١٠٢. Carnegie Mellon University, the Robotics nstitute Pittsburg, ٢٠٠٠.
- [١٠] Gary Bishop and Greg Welch. An introduction to the kalman filter. *Proc of SIGGRAPH*, Course, ٨:٢٧٥٩٩-٣١٧٥, ٢٠٠١.
- [١١] Chris Stauffer and W Eric L Grimson. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition*, ١٩٩٩. IEEE Computer Society Conference on., volume ٢. IEEE, ١٩٩٩.
- [١٢] S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler. ٢٠٠٠. Tracking groups of people. *Comput. Vis. Image Understand.* ٨٠, ١ (Oct. ٢٠٠٠), ٤٢-٥٦.
- [١٣] Cheng, F. & Chen, Y. (٢٠٠٦). Real time multiple objects tracking and identification based on discrete wavelet transform, *Journal of the pattern Recognition Society*, pp ١١٢٦-١١٣٩.
- [١٤] Czyz, J.; Ristic, B. & Macq, B. (٢٠٠٧). A particle filter for joint detection and tracking of color objects", *Image and Vision Computing*, Vol. ٢٥, No. ٨, pp ١٢٧١-١٢٨١.
- [١٥] Dong Kwon Park, Ho Seok Yoon and Chee Sun Won, "Fast Object Tracking In Digital Video," *IEEE Trans. Consumer Electronics*, vol. ٤٦, no. ٣, pp. ٧٨٥-٧٩٠, Aug. ٢٠٠٠.
- [١٦] Yining Deng, and B.S. Manjunath, "Unsupervised Segmentation of Color-Texture Regions in Images and Video," *IEEE Trans. Pattern Analysis And Machine Intelligence*, vol. ٢٣, no. ٨, pp. ٨٠٠-٨١٠, Aug. ٢٠٠١.

- [17] Yaakov Tsaig and Amir Averbuch, "Automatic Segmentation of Moving Objects in Video Sequences: A Region Labeling Approach," IEEE Trans. Circuits And Systems For Video Technology, vol. 12, no. 7, pp. 697-712, Jul. 2002.
- [18] R. Venkatesh Babu, K. R. Ramakrishnan, Member, IEEE, and S. H. Srinivasan, "Video Object Segmentation: A Compressed Domain Approach," IEEE Trans. Circuits And Systems For Video Technology, vol. 13, no. 4, pp. 462-474, Apr. 2003.
- [19] Vasileios Mezaris, Ioannis Kompatsiaris, and Michael G. Strintzis "Video Object Segmentation Using Bayes-Based Temporal Tracking and Trajectory- Based Region Merging," IEEE Trans. Circuits And Systems For Video Technology, vol. 13, no. 6, pp. 782-790 Jun. 2003.
- [20] Weiming Hu, Xi Li, Wenhan Luo, Xiaoqin Zhang, Stephen Maybank, and Zhongfei Zhang, "Single and Multiple Object Tracking Using Log-Euclidean Riemannian Subspace and Block-Division Appearance Model," IEEE Trans. Pattern analysis and machine intelligence, vol. 34, no. 12, pp. 2320-2330, Dec. 2012.
- [21] Chang Huang, Yuan Li, and Ramakant Nevatia Fellow, IEEE, "Multiple Target Tracking by Learning-Based Hierarchical Association of Detection Responses," IEEE Trans. Pattern analysis and machine intelligence, vol. 30, no. 4, pp. 898-910, Apr. 2008.
- [22] Rana Farah, J. M. Pierre Langlois, and Guillaume-Alexandre Bilodeau, "Catching a Rat by Its Edglets," IEEE Trans. Image Process., vol. 22, no. 2, pp. 668-678, Feb. 2013.
- [23] Shao-Yi Chien, Wei-Kai Chan, Yu-Hsiang Tseng, and Hong-Yuh Chen, "Video Object Segmentation and Tracking Framework With Improved Threshold Decision and Diffusion Distance," IEEE Trans. Circuits and systems for video technology, vol. 23, no. 6, pp. 911-924, Jun. 2013.
- [24] YILMAZ, A., JAVED, O., and SHAH, M. 2006. Object Tracking: A survey. ACM Comput. Surv. 38, 4, Article 13, 40 pages
- [25] Haibin Ling David W. Jacobs, "Shape Classification Using the Inner Distance." IEEE Transactions on Pattern Analysis and Machine Intelligence, 286-299, 2007
- [26] K.Srinivasan, K.Porkumaran, G.Sainarayanan, "Improved Background Subtraction Techniques For Security In Video Applications" 3rd international conference on Anti-Counterfeiting, security, and identification in communication, IEEE ,pp 114-117 - 2009.
- [27] Barga Deori and Dalton Meitei Thounaojam "A Survey On Moving Object Tracking In Video," International Journal on Information Theory (IJIT), vol.3, no.3, pp. 31-46 Jul. 2014.
- [28] Ritika and Gianetan Singh Sekhon, "Moving Object Analysis Techniques In Videos - A Review," IOSR Journal of Computer Engineering, vol. 1, Issue 2, pp. 07-12, May-June 2012.
- [29] Primet, M., Moisan, L.: Point tracking: an a-contrario approach. Preprint MAP0 no. 2012-06 (2012)
- [30] <http://en.wikipedia.org/>
- [31] S.-C. S. Cheung and C. Kamath "Robust techniques for background subtraction in urban traffic video", Visual Communications and Image Processing, 2004, vol. 0308, no. 1, (2004), pp. 881-892.
- [32] Himani S. Parekh, Darshak G. Thakore , Udesang K. Jaliya, "A Survey on Object Detection and Tracking Methods," International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 2, pp. 2970-2978, Feb. 2014.

- [33] ARULUMPALAM, M. ,MASKEL ,S. ,GORDON, N. and CLAPP ,T. A Tutorial on Particle Filters for Online Nonlinear/non-Gaussian Bayesian Tracking. 2002. IEEE Trans on Signal Processing, vol. 50,no.2 ,pp 178-188
- [34] ISARD ,M. and BLAKE ,A., Condensation-conditional density propagation for visual tracking. International journal of computer vision, 29 ,pp20-28 1998.
- [35] CUEVAS, E. , ZALDIVAR ,D .and ROJAS ,R. Kalman Filter for Vision Tracking. Technical Report B 00-122000.
- [36] COMANICIU, D., RAMESH, V. AND MERR, P. Kernel based-object tracking. IEEE Trans. Patt. Analy. Mach. Intell.20, 068-070. 2003.
- [37] Longyin Wen,Wenbo Li,Junjie Yan,Zhen Lei,Dong Yi,Stan Z. Li,” Multiple Target Tracking Based on Undirected Hierarchical Relation Hypergraph”, IEEE Conference on Computer Vision and Pattern Recognition IEEE,pp 1282-1289,2012
- [38] T. Yang, S. Z. Li, Q. Pan, and J. Li. Real-time multiple objects tracking with occlusion handling in dynamic scenes. In CVPR, pp 970-970, 2000.
- [39] H. Jiang, S. Fels, and J. J. Little. A linear programming approach for multiple object tracking. In CVPR, 2007.
- [40] L. Zhang, Y. Li, and R. Nevatia. Global data association for multiobject tracking using network flows. In CVPR, 2008.
- [41] C.-H. Kuo, C. Huang, and R. Nevatia. Multi-target tracking by online learned discriminative appearance models. In CVPR, pp 680-692, 2010.
- [42] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Globally-optimal greedy algorithms for tracking a variable number of objects. In CVPR, pp 1201-1208, 2011.
- [43] W. Brendel, M. R. Amer, and S. Todorovic. Multiobject tracking as maximum weight independent set. In CVPR, pp 1273-1280, 2011.
- [44] A. R. Zamir, A. Dehghan, and M. Shah. GMCP-tracker: Global multi-object tracking using generalized minimum clique graphs. In ECCV, pp 343-356, 2012.

Abstract:

Multi-target tracking is an interesting but challenging task in computer vision field. Most previous data association based methods merely consider the relationships (e.g. appearance and motion pattern similarities) between detections in local limited temporal domain, leading to their difficulties in handling long-term occlusion and distinguishing the spatially close targets with similar appearance in crowded scenes. In this paper, a novel data association approach based on undirected hierarchical relation hypergraph is proposed, which formulates the tracking task as a hierarchical dense neighborhoods searching problem on the dynamically constructed undirected affinity graph. The relationships between different detections across the spatiotemporal domain are considered in a high-order way, which makes the tracker robust to the spatially close targets with similar appearance. Meanwhile, the hierarchical design of the optimization process fuels our tracker to long-term occlusion with more robustness. Extensive experiments on various challenging datasets (i.e. ParkingLot), including both low and high density sequences, demonstrate that the proposed method performs favorably against the state-of-the-art methods.



Computer Engineering Faculty
Artificial intelligence

**A Thesis Submitted In Partial Fulfillment for the Degree of M.Sc in
Computer Engineering**

Multiple Target Tracking Based on Undirected Hierarchical Relation Hypergraph

By
Hamid Alimohamadi

Supervisor
Aliakbar Pouyan

Advisor
Morteza Zahedi

September 2015