

ردیابی بی‌درنگ اشیای متحرک در محیط‌های شلوغ با تحلیل تصاویر دوبعدی تک‌دوربینی با استفاده از یک فیلتر ذره‌ای توسعه‌یافته غنی‌شده

نصرالله مقدم چرکری^{۱*}، حمیدرضا شایق بروجنی^۲

۱. استادیار مهندسی کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه تربیت مدرس

۲. دانشجوی دکترا نرم‌افزار، دانشکده برق و کامپیوتر، دانشگاه تربیت مدرس

moghadam@modares.ac.ir

تاریخ پذیرش: ۱۳۹۲

تاریخ دریافت: ۱۳۸۹

چکیده

در این مقاله، روشی جدید برای ردیابی چند شیء متحرک در تصاویر ویدئویی ارائه شده است. با این روش، می‌توان بی‌درنگ اشیای متحرک در محیط‌های همراه با انسداد، هم‌پوشانی و تغییر جهت-های متعدد را ردیابی کرد. روش پیشنهادی، توسعه‌ای بر فرآیند فیلتر ذره‌ای است که در این مقاله با عنوان فیلتر ذره‌ای غنی‌شده ارائه می‌شود. در این روش، خصوصیتی از شیء متحرک که امکان به‌کارگیری مستقیم آن‌ها در فرآیند محاسبات احتمال پسین در فیلتر ذره‌ای وجود ندارد، به عنوان اطلاعات اضافی به فرآیند ردیابی افزوده می‌شوند. به این ترتیب، تخمین احتمالاتی فیلتر ذره‌ای در هر قاب پس از انطباق با مشاهدات، با استفاده از اطلاعات غنی‌کننده، تصحیح شده و تخمین مناسب‌تری از هر شیء قبلی در قاب جاری ایجاد می‌شود. ارزیابی روش پیشنهادی با مجموعه داده-هایی از جمله PETS09 و مقایسه نتایج آن با سه رویکرد دیگر مبتنی بر فیلتر ذره‌ای، نشان از بهبود دقت و کاهش خطای ردیابی به میزانی در حدود ۱۲٪ دارد.

واژگان کلیدی: تحلیل تصاویر ویدئویی، تصاویر ویدئویی با دوربین ثابت، ردیابی اشیای متحرک، فیلتر ذره‌ای، فیلتر ذره‌ای غنی‌شده.

۱. مقدمه

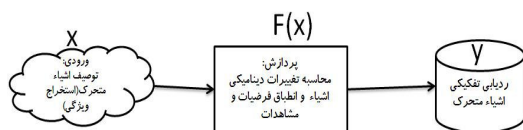
یکی از شاخه‌های پردازش تصویر که در سال‌های اخیر مورد توجه بسیاری قرار گرفته است، پردازش ویدئویی^۱ است. قوی‌تر شدن سخت‌افزارهای محاسباتی و رشد نیاز به وجود سامانه‌هایی که از این روش استفاده می‌کنند، مهم‌ترین دلایل این امر است [۱]. پردازش تصاویر ویدئویی برای تشخیص، تعقیب و تحلیل اشیای متحرک موجود در صحنه، یکی از پرکاربردترین زمینه‌های تحقیقاتی در این حوزه در دهه گذشته بوده است.

یکی از مهم‌ترین کاربردها در این حوزه، کاربردهای نظارتی^۲ است [۲]. نظارت خودکار مکان‌های عمومی نظیر سالن‌های مترو، بانک‌ها، فرودگاه‌ها، ورزشگاه‌ها و...، کنترل خودکار بیماران و فعالیت آن‌ها در بیمارستان‌ها و مراکز نگهداری کودکان و سالمندان، کنترل خودکار ترافیک و ثبت تخلفات رانندگی، مواردی از این دست است. در سال ۲۰۰۸ تنها در سالن‌های متروی لندن بیش از ۱۰۰۰۰ دوربین مداربسته وجود داشت [۳]. نظارت انسانی این تعداد دوربین، قطعاً فرآیندی پرهزینه و مشکل است.

ردیابی شیء متحرک از اساسی‌ترین بخش‌های فرآیند نظارت ویدئویی به‌شمار می‌رود. تعقیب تفکیکی اشیای متحرک در طول دنباله قاب‌های تصویر در یک ویدئو، به‌طوری‌که در اثر انسداد یا هم‌پوشانی با موانع و اشیای دیگر توسط سامانه ردیاب گم نشوند، ردیابی نامیده می‌شود [۱]. برای ردیابی اجسام متحرک، دو رویکرد مورد توجه قرار می‌گیرد: در رویکرد اول پس از مشاهده شیء، با مقایسه ویژگی‌های شیء مشاهده‌شده با اشیای قبلی در

تاریخچه مشاهدات، شناسه آن به‌دست می‌آید [۲]. در رویکرد جامع‌تر، که اغلب روش‌های ردیابی بر آن استوار است، از یک چارچوب احتمالاتی برای تخمین مکان یا حالت شیء استفاده‌شده و مشاهدات قاب جاری با فرضیات احتمالی محاسبه‌شده تطبیق داده می‌شوند [۱].

برای بهبود یک فرآیند ردیاب، که از چارچوب احتمالاتی پیروی می‌کند، می‌توان به صورت بالا به پایین عمل کرد و به بهبود چگونگی استخراج ویژگی‌های شیء متحرک پرداخت و یا به صورت پایین به بالا به مسئله نگاه کرد و چگونگی تولید فرضیه تغییرپویایی را بهبود بخشید [۴]. درواقع چنانچه مطابق شکل (۱) به سامانه ردیاب به عنوان یک تابع f با ورودی x و خروجی غیرقطعی y نگاه کنیم، بهبود خروجی غیرقطعی y هم با بهبود ورودی x قابل انجام است و هم با بهبود فرآیند عملکردی تابع f .



شکل (۱). بهبود خروجی y در فرآیند ردیابی هم با تغییر یا بهبود تابع f امکان‌پذیر است و هم با بهبود چگونگی توصیف شیء یا همان ورودی تابع (x)

یکی از بهترین فرآیندهای احتمالاتی ردیابی که از تطابق احتمالات پسین و مشاهدات فعلی برای ردیابی استفاده می‌کند، چارچوب فیلتر ذره‌ای است [۵]. از چالش‌های حل‌نشده در این روش هنگامی است که چندین شیء متحرک در صحنه علاوه بر مشابهت‌های ظاهری، حرکات تصادفی و غیرقابل پیش‌بینی دارند. در این مواقع، تخمین احتمالاتی پسین ویژگی‌ها یا مکان هر شیء، که بر اساس تاریخچه مشاهدات پیشین انجام می‌شود، ممکن است گمراه‌کننده باشد. ایده اصلی مقاله،

1. Video processing.

2. Surveillance applications.

استفاده از داده‌هایی مستقل از ورودی فیلتر ذره‌ای و اعمال مستقیم آن به ورودی تابع ردیاب با هدف بهبود فرآیند تصمیم و اصلاح آن است. این داده‌ها به عنوان اطلاعات غنی‌کننده، اطلاعات بیشتری را به فرآیند تصمیم‌گیری در مورد انطباق اشیای متحرک قاب جاری و قاب‌های قبلی با هدف ردیابی می‌افزایند. چنین نگرشی را فیلتر ذره‌ای غنی‌شده می‌نامیم که برای اولین بار در این مقاله به آن توجه شده است.

در ادامه، در بخش دوم، مهم‌ترین کارهای انجام‌شده و رویکردهایی که در بخش‌های متفاوت فرآیند ردیابی به‌کار رفته است، معرفی می‌شود. در بخش سوم، فرضیات و محیط مسئله معرفی خواهد شد. در بخش چهارم در مورد چارچوب احتمالاتی سامانه ردیاب و چارچوب فیلتر ذره‌ای به عنوان اصلی‌ترین چارچوب احتمالاتی، بحث خواهد شد. بخش پنجم به ارائه روش پیشنهادی، ایده اولیه و فرآیند کار اختصاص دارد. نتایج آزمایش‌ها در بخش شش و جمع‌بندی در بخش هفتم خواهد بود.

۲. مروری بر رویکردها و مهم‌ترین کارهای انجام‌شده

اغلب روش‌های ردیابی از یک چارچوب احتمالاتی^۱ برای پیش‌بینی احتمال پسین حالت شیء در قاب^۲ جاری و قاب‌های بعدی بر اساس مشاهدات در قاب‌های قبلی استفاده می‌کنند [۱]. کارهای فراوانی در این حوزه مطرح شده است که از جنبه‌های مختلف قابل بررسی است. یکی از این جنبه‌ها، تفاوت در مدل مشاهده است [۲].

برخی از روش‌ها مبتنی بر ناحیه شیء متحرک است

[۶-۱۶]. در [۶] از هیستوگرام رنگ اشیای متحرک برای تشکیل حباب‌های حرکتی استفاده می‌شود. در [۷ و ۸] از مدل رنگی پوست برای ردیابی، بهره گرفته شده است. در [۹] از بافت با استفاده از توصیف‌کننده LBP برای تشخیص و تعقیب شیء متحرک استفاده شده است. در [۱۰ و ۱۱] از ویژگی‌های ظاهری متفاوت برای شناسایی عابران در کاربرد رانندگی خودکار و شناسایی مرز سر افراد استفاده شده است. از ویژگی‌های لبه به عنوان هیستوگرام راستای لبه نیز در [۱۲ و ۱۳] استفاده شده است. به‌طور کلی، ویژگی‌های لبه به علت مقاومت در برابر تغییرات روشنایی، کارایی مناسب دارند. در [۱۴ و ۱۵] از ویژگی گوشه‌ها به عنوان ویژگی متمایزکننده استفاده شده است. در [۱۶] از یک مدل سه‌بعدی تولیدشده با استفاده از دو دوربین برای توصیف حالت راه‌رفتن افراد استفاده می‌شود. در [۳۲] از ضرایب موجک برای ردیابی در محیط‌های شلوغ استفاده شده است.

علاوه بر ویژگی‌های ظاهری، ویژگی‌های حرکتی نیز به‌طور گسترده در تولید مدل مشاهده به‌کار گرفته می‌شود. در [۱۷] از محاسبه سرعت شیء با استفاده از میزان جابه‌جایی شیء برای تشخیص شیء استفاده شده است. جهت حرکت شیء نیز در برخی کاربردها [۱۸]، که اشیاء الگوهای تقریباً ثابت حرکتی دارند، به عنوان ویژگی در نظر گرفته شده است. تغییرات اندازه شیء متحرک در طول زمان نیز معیار خوبی برای تمایز اشیای متحرک از یکدیگر است [۱۷ و ۱۹]. در [۲۰]، پیشنهادهایی به منظور استفاده از ویژگی‌های سایه در محیط روشن مطرح شده است.

به‌طور کلی ویژگی‌های بیانگر پویایی شیء متحرک از جمله ویژگی‌های حرکتی شیء در مدل‌های احتمالاتی ردیابی برای تخمین ویژگی‌های شیء در هر قاب جاری

1. Probabilistic framework.

2. Frame.

استفاده شده است. در [۲۹] از یک فیلتر ذره‌ای به همراه یک مرحله خوشه‌یابی برای ردیابی چند شیء استفاده شده است. در [۳۰] از یک مدل گرافیکی مبتنی بر فیلتر ذره‌ای برای ترکیب ویژگی‌ها استفاده شده است. در این مدل ویژگی‌های رنگ و شکل به صورت مکمل به کار رفته‌اند. یکی از موارد مهم در فیلتر ذره‌ای، الگوریتم انتشار تابع چگالی احتمال شرطی است که در [۳۱ و ۳۳] روش‌هایی برای بهبود آن پیشنهاد شده است. در [۳۳] از یک الگوریتم نمونه‌برداری خلاصه‌سازی شده توسعه یافته برای نمونه‌برداری‌های مجدد و وزن‌دار کردن ذرات استفاده شده است. فیلترهای ذره‌ای سه‌بعدی که از ترکیب دو مرحله‌ای ویژگی‌ها برای مدل کردن تابع چگالی احتمال پسین استفاده می‌کنند، از جمله جدیدترین روش‌های ردیابی محسوب می‌شوند [۳۴]. این فیلترها اغلب برای محیط‌هایی که با چند دوربین، تصویربرداری شده‌اند کاربرد دارند. روش‌های مبتنی بر یادگیری ماشین به کارگیری روش‌های دسته‌بندی نظارتی و نیمه نظارتی نیز برای ردیابی اشیای متحرک در چند مرجع به کار گرفته شده‌اند [۳۵ و ۳۶]. رویکردهای متنوع مبتنی بر فیلتر ذره، با وجود مقاومت در برابر انسدادها و هم‌پوشانی‌های جزئی و موقتی، برای ردیابی چند شیء متحرک در یک محیط شلوغ که اشیاء، ویژگی‌های ظاهری تاحدی مشابه و ویژگی‌های حرکتی تصادفی داشته باشند، مناسب نیست. علت اصلی این امر، ناتوانی در به کارگیری تمامی ویژگی‌های قابل استخراج از شیء متحرک در فیلتر ذره‌ای به دلیل افزایش پیچیدگی محاسباتی و عدم کارایی برخی ویژگی‌ها در محاسبه احتمالات شرطی پسین است.

به عنوان مدل مشاهده، کارآمدتر است؛ اما استخراج آن‌ها نسبت به ویژگی‌های ظاهری پیچیده‌تر است. در اغلب روش‌ها، از جمله روش پیشنهادی در این مقاله، از ترکیب مناسبی از ویژگی‌های ظاهری و حرکتی با توجه به کاربرد فرآیند ردیابی استفاده می‌شود.

کارهای انجام شده در فرآیند ردیابی را همچنین می‌توان بر اساس مدل احتمالاتی به کار گرفته شده و چگونگی تحلیل مدل مشاهده اشیای متحرک، بررسی کرد [۲]. یکی از مدل‌های پایه موجود، فیلتر کالمن است. فیلتر کالمن^۱ یک تابع توزیع گاوسی از حالات هدف را به دست می‌دهد [۲۱]. به علت فرض خطی و گاوسی بودن حالت هدف در فیلتر کالمن استفاده از آن در بسیاری از محیط‌ها ممکن نیست (محیط‌های شلوغ با حرکات تصادفی و غیرثابت اشیاء) و بنابراین از توسعه‌هایی از این فیلتر در بسیاری از کارها استفاده شده است [۲۲]. در [۲۳] از فیلتر کالمن توسعه یافته با ترکیبی از گاوسی‌ها برای تخمین حالت بعدی شیء، که از ترکیبی از ویژگی‌های هندسی، حرکت و رنگ به دست می‌آید، استفاده شده است. [۲۴ و ۲۵] از معیار و ضریب فاصله برجستگی استفاده کرده‌اند. در [۲۶] ردیابی حباب با استفاده از تابع هزینه به کار گرفته شده است. یکی از بهترین و جدیدترین روش‌های ردیابی، ردیابی در چارچوب فیلتر ذره‌ای است [۱ و ۲۷]. در این روش، تابع چگالی احتمال پسین حالت شیء توسط تعدادی ذره با وزن‌های متناوب مدل می‌شود. این روش، توانایی ردیابی در حالات غیرخطی و غیرگاوسی را به خوبی داشته و از این لحاظ بر فیلتر کالمن برتری دارد. در [۲۸] از یک چارچوب فیلتر ذره‌ای با فرض مستقل بودن ویژگی‌های رنگ، حرکت و صدا

1. Kalman filter.

۳. فرضیات محیط ردیابی و فضای مسئله

رویکرد موردنظر در بخش‌های مختلف فرآیند ردیابی و ارائه راهکار برای بهبود آن‌ها، به نوع فضای مسئله، محیط مورد ردیابی، چگونگی تصویربرداری آن و مواردی از این دست وابسته است.

در این مقاله، هدف، ردیابی اشیای متحرک در محیطی با خصوصیات و نوع تصویربرداری زیر است:

- فضای مورد ردیابی وسعتی دارد که می‌توان با یک دوربین، کل آن فضا را پوشش داد؛ بنابراین تمام صحنه‌های قابل تحلیل، دوبعدی است.

- دوربین به صورت کاملاً ثابت از محیط، تصویربرداری می‌کند.

- فضای مسئله هم شامل تصویربرداری‌های درونی و هم شامل تصویربرداری‌های بیرونی است.

- اجسام متحرک در صحنه، انسان هستند.

- حرکت اشیا در صحنه از هیچ قانون خاصی (سرعت یا جهت ثابت یا متناوب، شتاب ثابت و...) پیروی نمی‌کند.

- روشنایی محیط، ممکن است دچار تغییرات شود.

- تعداد اشیای متحرک در هر قاب، حداقل ۰ و حداکثر ۱۵ است.

- اشیای متحرک، تشابهات ظاهری و حرکتی زیادی با هم دارند.

- اجسام متحرک در طول دنباله تصاویر به چند شیء مجزا تبدیل نشده و چند شیء مجزا نیز، جز در به هم پیوستن موقتی اجسام، به یکدیگر متصل نمی‌شوند.

در این برای انجام آزمایش‌ها از سه مجموعه داده به شرح زیر استفاده شده است:

- مجموعه استاندارد [۳۷PETS06 IEEE]: این

مجموعه شامل ۷ تصویربرداری (زیرمجموعه) است که

توسط ۴ دوربین، فیلمبرداری شده است و مربوط به جاگذاشتن یک کیف در تصویربرداری‌هایی است که به تدریج پیچیده‌تر می‌شود.

- مجموعه استاندارد [۳۸PETS09 IEEE]:

جدیدترین مجموعه داده در این زمینه است که با ۸ دوربین در محیط بیرونی در فضای سبز دانشگاه Utah فیلمبرداری شده است. تصویربرداری‌های حرکت فردی و گروهی افراد، دوزدن‌های تصادفی و... در این مجموعه به چشم می‌خورد.

- مجموعه‌ای ترکیبی از برخی داده‌های دانشگاه کمبریج که شامل ویدئوهایی از حرکت اتومبیل‌ها در خیابان و حرکت انسان‌ها در سالن متروست.

در این مقاله از تصاویر هر دوربین به صورت مستقل استفاده شده و هم‌پوشانی زمانی مکانی تصاویر از دوربین‌های مختلف مدنظر نیست.

نمونه‌هایی از قاب‌های دنباله تصاویر از مجموعه داده‌های نام‌برده‌شده در شکل (۲) نشان داده شده است.

۴. چارچوب احتمالاتی سامانه ردیاب

یک سامانه ردیاب احتمالاتی، ۴ بخش دارد [۴]. دو بخش اول به چگونگی مشاهده شیء و دو بخش دوم به چگونگی تفسیر شیء بر اساس احتمالات حالت شیء بر اساس گذشته مشاهدات آن مربوط است. این بخش‌ها عبارت‌اند از:

نمایش هدف^۱: شناسایی اهداف متحرک موردنظر.

نمایش مشاهده^۲: توصیف اهداف متحرک شناخته‌شده (استخراج ویژگی آن‌ها).

1.Target Representation.

2.Observation Representation.

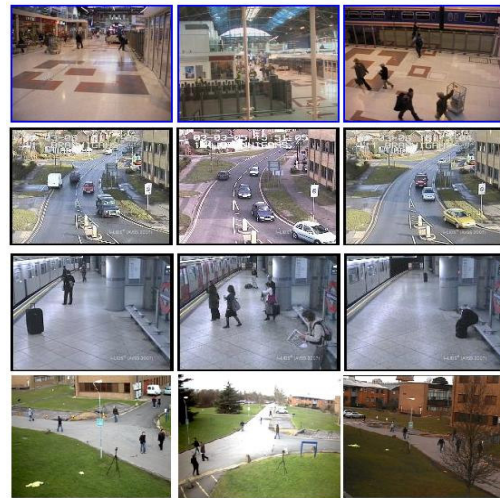
روابط سیستم به فرم احتمالاتی و به روز کردن اطلاعات حالت سیستم بر مبنای دریافت اندازه گیری های جدید، ما را به سمت روش تخمین مبتنی بر قانون بیز^۳ رهنمون می کند. بر اساس آن می توان یک چارچوب عمومی برای مسائل تخمین پویای^۴ حالت شیء ارائه کرد.

در روش تخمین پویای حالت مبتنی بر قانده بیز، هدف به دست آوردن تابع چگالی احتمال پسین^۵ حالت سیستم بر مبنای تمامی اطلاعات موجود (شامل مجموعه تمام اندازه گیری ها) است. این تابع چگالی احتمال یک راه حل کامل مسئله تخمین حالت سیستم است. در بسیاری از مسائل به یک تخمین جدید از شیء یا سیستم به ازای هر اندازه گیری جدیدی که به دست می آید، نیاز می شود. در این حالت یک فیلتر برگشتی^۶ راه حلی مناسب خواهد بود [۲۸]. روش فیلتر بازگشتی بدین معنی است که داده دریافتی را می توان به طور متوالی و نه به یکباره پردازش کرد؛ بنابراین نیازی به نگهداری مجموعه کامل داده ها (اندازه گیری ها) یا پردازش مجدد داده های موجود در صورت دریافت اندازه گیری جدید، نخواهد بود. چنین فیلتری از دو مرحله پیش بینی^۷ و به روز رسانی^۸ تشکیل می شود. در مرحله پیش بینی از مدل سیستم برای پیش بینی تابع چگالی احتمال حالت سیستم یا شیء در لحظه آینده بر اساس اندازه گیری ها تا لحظه جاری استفاده می شود. در مرحله به روز رسانی از آخرین اندازه گیری برای بهبود تابع

تولید فرضیه^۱: محاسبه فرضیه تغییرات پویای حالت مدل مشاهده شیء.

اندازه گیری فرضیه^۲: تطبیق نمایش مشاهده با فرضیه تولید شده.

در یک سیستم ردیاب، حالت های هدف با $\mathbf{X}_t$ و مشاهده با $\mathbf{Z}_t$ نمایش داده می شوند. گذشته حالت های هدف با $\mathbf{x}_t = (\mathbf{x}_1, \dots, \mathbf{x}_t)$ و گذشته مشاهدات با $\mathbf{z}_t = (\mathbf{z}_1, \dots, \mathbf{z}_t)$ نمایش داده می شوند. در حقیقت بردارهای فوق، نمایش فضای حالت، حالت هدف و مشاهدات است. هدف این چارچوب احتمالاتی در واقع استفاده از این بردارها در شناخت مناسب مشاهده ها در زمان t و تطبیق مناسب آن ها با حالت های موجود در زمان $t-1$ است.



شکل (۲). نمونه هایی از مجموعه دنباله تصاویر استاندارد مورد آزمایش در پژوهش. تمام این ویدئوها با یک دوربین ثابت تهیه شده و صحنه های انتخابی، شامل حداکثر ۱۵ شیء متحرک است. تمام اشیای متحرک، انسان یا وسیله نقلیه است.

3. Bayesian approach.

4. Dynamic.

5. Posterior Probability Density Function.

6. Recursive Filter.

7. Prediction.

8. Updating.

1. Hypotheses generation.

2. Hypotheses measurement.

چگالی احتمال حاصل از پیش‌بینی استفاده می‌شود. در این مرحله از قضیه بیز^۱، استفاده می‌شود.

همان‌طور که ذکر شد، هدف، به‌دست‌آوردن تابع چگالی احتمال پسین حالت شی $\mathbf{x}_k$ بر اساس اندازه‌گیری‌ها تا لحظه K ام، $\mathbf{z}_{1:k}$ ، است. تابع چگالی احتمال پیشین اولیه، $p(\mathbf{x}_0 | \mathbf{z}_0)$ ، دانسته فرض می‌شود. فرض کنید که تابع چگالی احتمال $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ در لحظه $k-1$ موجود باشد. در مرحله پیش‌بینی، تابع چگالی احتمال پیشین حالت شیء به کمک معادله چپمن-کولوموگروف^۲ به‌دست می‌آید.

(۱)

$$P(x_k | x_{k-1}, z_{1:k-1}) = \int P(x_k | x_{k-1}) P(x_{k-1} | x_{k-2}, z_{1:k-2}) dx_{k-1}$$

در رابطه (۱) از این حقیقت که $P(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_{1:k-1}) = P(\mathbf{x}_k | \mathbf{x}_{k-1})$ استفاده می‌شود. در لحظه k ، اندازه‌گیری $\mathbf{z}_k$ به‌دست می‌آید که از آن در به‌روزرسانی تابع چگالی احتمال پیشین توسط قاعده بیز استفاده می‌شود.

$$P(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{P(\mathbf{z}_k | \mathbf{x}_k) P(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{P(\mathbf{z}_k | \mathbf{z}_{1:k-1})} \quad (2)$$

۴-۱. چارچوب فیلتر ذره‌ای

چارچوب فیلتر ذره‌ای در مواردی که تابع چگالی احتمال پسین حالت سیستم دارای چند ماکزیمم بوده و گوسی نباشد، اغلب با استفاده از الگوریتم نمونه‌برداری

فاکتور شده به‌کار می‌رود [۲۸]. این روش از ۴ بخش تشکیل می‌شود:

تولید ذرات محاسبه احتمال پسین حالت ذرات نمونه-برداري مجدد ذرات در قاب جاری تطبیق ذرات جدید (مشاهدات) با حالت‌های پسین ذرات قبلی (فرضیه‌ها) و اصلاح وزن ذرات این فرآیند به‌صورت بازگشتی انجام می‌شود.

چارچوب فیلتر ذره‌ای، یک چارچوب کلی است که الگوریتم‌های متفاوتی برای پیاده‌سازی آن وجود دارد؛ اما الگوریتمی که بیش از همه در بینایی ماشین استفاده می‌شود، الگوریتم انتشار تابع چگالی احتمال شرطی^۳ است. ما از همین الگوریتم در روش پیشنهادی استفاده کرده و آن را بهبود داده‌ایم.

اگر توزیع شرطی بر مبنای حالت سیستم در لحظه قبل به‌عنوان تابع اهمیت انتخاب شود.

$$p(x_k^{(i)} | x_{k-1}^{(i)}, z_{1:k}) = p(x_k^{(i)} | x_{k-1}^{(i)}) \quad (3)$$

معادله به‌روزرسانی وزن‌های ذرات به‌صورت زیر تبدیل می‌شود.

$$w_k^{(i)} \propto p(z_k | x_k^{(i)}) \quad (4)$$

برای انجام الگوریتم انتشار تابع چگالی شرطی مراحل زیر در هر گام انجام می‌گیرد. در هر لحظه k از مجموعه $\{x_{k-1}^{(i)}, w_{k-1}^{(i)}, c_{k-1}^{(i)}\} \quad i = 1 \dots N$ که شامل ذرات، وزن-های مربوطه و توزیع تجمعی آن‌ها، C_{k-1}^i ، است، مجموعه جدیدی برای لحظه جاری، k ،

$$\{x_k^{(i)}, w_k^{(i)}, c_k^{(i)}\} \quad i = 1 \dots N$$

به‌دست می‌آید:

نمونه‌های $x_k^{(i)}$ به ترتیب زیر تولید می‌شوند:

3. Conditional Density Propagation Algorithm (CONDENSATION).

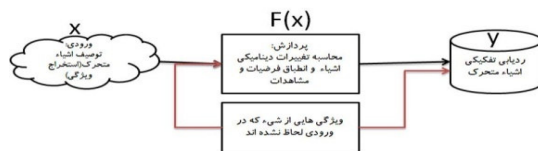
1. Bayesian Theorem.

2. Chapman-kolmogorov.

همچنین با وجود کارایی سامانه‌های ردیاب بر اساس فیلتر ذره‌ای، این روش‌ها در برابر انسدادهای جزئی و کلی در هنگامی که دو شیء، ویژگی‌های مشابهی داشته باشند، حساس است.

ایده‌ای که در این مقاله برای بهبود روش‌های مبتنی بر فیلتر ذره‌ای موجود استفاده شده، بر دو مورد استوار است: ۱. بهبود ویژگی انتخابی مورد ردیابی با ترکیب مناسب ویژگی‌ها با ضرایب مناسب.

۲. استفاده از یک دانش افزوده در کنار تصمیم خروجی سیستم ردیاب برای تأیید یا رد نتیجه و اصلاح آن. شکل (۳) دیدی کلی از روش پیشنهادی را در بهبود فرآیند ردیابی نشان می‌دهد. درواقع در روش پیشنهادی به بهبود ورودی تابع ردیاب با انتخاب وزن‌دار ویژگی‌های ظاهری و حرکتی و همچنین به بهبود فرآیند تصمیم‌گیری تابع ردیاب با افزودن اطلاعاتی به عنوان اطلاعات غنی-کننده به موازات فرآیند قبلی می‌پردازیم که نتایج قابل قبولی از آن به‌دست آورده‌ایم.



شکل (۳). ایده اساسی روش پیشنهادی، استخراج برخی ویژگی-ها اضافه بر ویژگی‌های مورد استفاده در فیلتر ذره‌ای و استفاده از آن‌ها به عنوان اطلاعات غنی‌کننده برای بهبود تابع ردیاب است.

۵-۱. ایده اولیه ساخت فیلتر ذره‌ای غنی-

شده در دو بعد

حتماً تاکنون به سینماهای موسوم به سه‌بعدی یا چهاربعدی رفته‌اید. سؤالی که در ذهن هر کس قبل از اولین حضور در یکی از این سینماها ممکن است پیش بیاید، چگونگی مشاهده سه و یا چهار بعد توسط چشم و

الف) یک عدد تصادفی r با توزیع احتمالی یک نواخت در بازه $[0, 1]$ تولید می‌شود.

ب) با تقسیمات باینری کوچک‌ترین j که $C_{k-1}^j \geq r$ باشد، یافت می‌شود.

ج) $x_k^{(i)}$ برابر $x_k^{(j)}$ قرار داده می‌شود.

د) با استفاده از مدل دینامیکی یا تابع چگالی احتمال پیشین، نمونه‌های جدید به‌صورت زیر تولید می‌شود:

$$x_k^i \sim p(x_k | x_{k-1} = x_{k-1}^{(i)}) \quad (5)$$

ه) وزن مربوط به نمونه جدید یا ذره جدید با استفاده از تابع شباهت یا تابع چگالی احتمال مشاهده به شرط حالت سیستم بر حسب مشاهده اندازه‌گیری شده به‌دست می‌آید.

$$w_k^i = p(z_k | x_k = x_k^{(i)}) = L(z_k | x_k = x_k^i) \quad (6)$$

آن‌گاه وزن‌های اندازه‌گیری شده با شرط $\sum_{i=1}^N w_k^i = 1$ نرمالیزه می‌شوند. و توزیع تجمعی آن‌ها که از روابط زیر به‌دست می‌آید، ذخیره می‌شود.

$$C_k^{(0)} = 0 \quad (7)$$

$$C_k^{(i)} = C_k^{(i-1)} + w_k^{(i)} \quad i = 1, \dots, N$$

آن‌گاه تخمین هر تابع، f ، از حالت سیستم x به-صورت زیر محاسبه می‌شود.

$$E[s(x_k)] = \sum_{i=1}^N w_k^{(i)} f(x_k^{(i)}) \quad (8)$$

اگر $f(x)=x$ انتخاب شود، تخمین فوق برابر میانگین X در لحظه k خواهد بود.

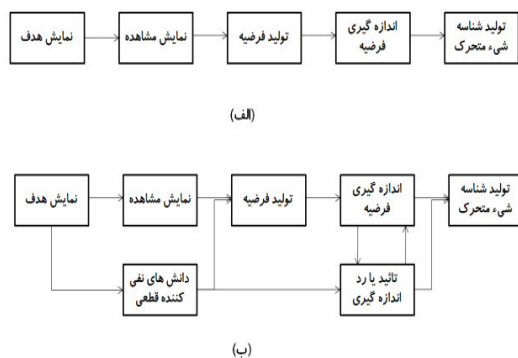
۵. توسعه فیلتر ذره‌ای با افزودن بعد سو

اگرچه فیلتر ذره‌ای روشی مقاوم و قوی در ردیابی است، قدرت آن به ویژگی انتخابی مورد ردیابی و همچنین تابع شباهت به‌دست آمده بر اساس این ویژگی بستگی دارد.

زبان، حاکی از ملیت بریتانیایی فرد دارد؛ اما با توجه به این که فرد هیچ گونه آشنایی با زبان انگلیسی ندارد، با قطعیت بالا می توان گفت که این فرد چنین ملیتی ندارد درحالی که نمی توان گفت چه ملیتی دارد.

چنانچه بتوان سؤالات نفی کننده مناسبی پیدا کرد، می توان به راحتی از تعلق اشتباه یک حدس به یک موجودیت جلوگیری کرد و این آن چیزی است که به عنوان بعد مجازی سوم به دنبال آن هستیم تا درک خود را از تعلق یک مشاهده به یک فرضیه اندازه گیری کامل کنیم. ایده دوم در روش پیشنهادی که باعث کارایی بالاتر آن در ردیابی شده است، به کارگیری وزن دار ویژگی های ظاهری و حرکتی مناسب به عنوان ورودی سامانه است که با استفاده از آزمایش های گوناگون بر روی تنوعی از ویژگی هایی که در مراجع مختلف از آن ها استفاده شده است، به دست آمده اند.

شکل (۴) چارچوب احتمالاتی ردیابی فعلی و چارچوب احتمالاتی پیشنهادی با افزودن اطلاعات غنی کننده به منظور تولید نتیجه بهتر را در کنار هم نشان می دهد.



شکل (۴). مقایسه انتزاعی فرآیند روش پیشنهادی (ب) و روش معمول ردیابی در فیلتر ذره ای (الف) تفاوت اساسی در بخش دانش های نفی کننده به عنوان بعد سوم غنی کننده و تأثیر آن در تخمین مشاهده است.

از یک تصویر ویدئویی دوبعدی است. در این سینماها، منظور از بعد سوم و چهارم، ابعادی خارج از تصویر دوبعدی است که به کامل کردن درک تصویر دوبعدی کمک می کند. حس های لامسه و بویایی و تکان خوردن صندلی بیننده مطابق با اتفاقی که در تصویر در حال رخ دادن است، ابعادی است که اغلب به عنوان بعدهای سوم و چهارم به کار رفته، درک فرد را از نمایش کامل تر و تصاویر را واقعی تر می کند.

ایده اصلی و اولیه ای که از آن در بهبود نگرش به یک تصویر دوبعدی و بهبود فرآیند مبتنی بر فیلتر ذره ای بهره گرفته ایم، استفاده از دانش هایی به عنوان بعد سوم و به عنوان اطلاعات غنی کننده، در کنار فرآیند تصمیم گیری فیلتر ذره ای به منظور بهبود درک شناسه اشیا است. مشکلی که در اینجا با آن روبه رو هستیم، نبود بعد دیگری به جز همان دو بعد تصویربرداری شده با یک دوربین است؛ بنابراین باید به نحوی بعد سوم را به صورت مجازی و از همان اطلاعات استخراجی از دو بعد، ایجاد کرد. برای این کار از ایده ویژگی نفی کننده بهره گرفته ایم. در بسیاری از مواقع، وجود یک ویژگی برای یک موجودیت در تصمیم گیری در مورد شناسه آن موجودیت به میزان بقیه ویژگی ها تأثیر دارد؛ اما نبود آن ویژگی برای یک موجودیت می تواند به صورت قطعی پاسخ نهایی به دست آمده از ترکیب ویژگی ها را نفی کند. این مطلب را با یک مثال شرح می دهیم:

فرض کنید هدف تشخیص ملیت یک فرد از روی ویژگی های مختلف آن است. یک فرد با پوست روشن، موهای بور، چشمان روشن، آشنا به زبان های فارسی و انگلیسی، احتمالاً یک فرد بریتانیایی است. حال، یک فرد با پوست روشن، موهای بور، چشمان روشن، آشنا به زبان فارسی و نه انگلیسی را تصور کنید. تمام ویژگی ها به جز

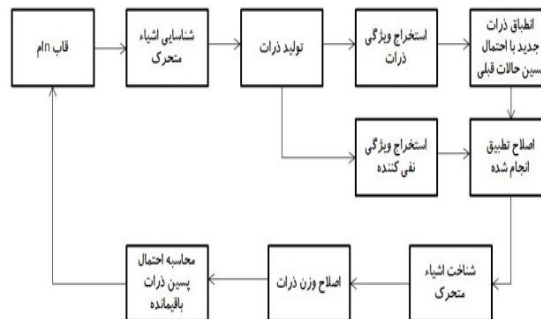
۱. استخراج ویژگی‌های مناسب برای فیلتر ذره‌ای و یافتن ویژگی‌های مناسب به عنوان دانش نفی‌کننده و وزن‌دهی آن‌ها.
۲. طراحی فیلتر ذره‌ای غنی‌شده با افزودن بخش دانش نفی‌کننده به عنوان بعد سوم غنی‌کننده.

۳-۵. استخراج ویژگی‌ها

تاکنون ویژگی‌های متنوعی در ردیابی اشیای متحرک مورد استفاده قرار گرفته است. برخی از آن‌ها ویژگی‌های ظاهری و برخی دیگر ویژگی‌های دینامیکی (حرکتی) شیء است. ویژگی‌های مناسب در جدول (۱) نشان داده شده است. به‌کارگیری و استخراج تمامی این ویژگی‌ها بسیار زمان‌بر بوده، پیچیدگی محاسباتی فرآیند ردیابی را افزایش می‌دهند؛ بنابراین لازم است تا بهترین ویژگی‌ها انتخاب شوند. همان‌طور که عنوان شد، در روش پیشنهادی به دو دسته ویژگی نیاز است. یک دسته ویژگی‌هایی که بیان‌کننده حالت شیء بوده و از ترکیب مناسب آن‌ها در فیلتر ذره‌ای برای محاسبه حالت پسین شیء استفاده می‌شود و دسته دیگر ویژگی‌هایی که پاسخی به سؤالات نفی‌کننده باشد.

ویژگی‌هایی که به عنوان دانش نفی‌کننده می‌تواند به-کار گرفته شود را می‌توان تا حدود زیادی از نوع ویژگی استخراجی شناخت؛ اما برای انتخاب سایر ویژگی‌ها در فیلتر ذره‌ای، احتیاج به آزمایش‌هایی است که کارایی ویژگی را معین می‌کند. در این مرحله، ویژگی‌های زیر برای استفاده به عنوان حالت شیء در فیلتر ذره‌ای کاندید است:

در روش پیشنهادی، هدف به‌کارگیری ویژگی‌هایی از اشیای متحرک دنباله تصاویر است که بتوان از آن‌ها به عنوان سؤالات نفی‌کننده برای اصلاح پاسخ فیلتر ذره‌ای استفاده کرد؛ زیرا به جز این اطلاعات، بعد و اطلاعات دیگری قابل دستیابی نیست.



شکل (۵). بلوک دیاگرام روش پیشنهادی برای ردیابی اشیای متحرک در قاب m

شکل (۵) مراحل روش پیشنهادی را نشان می‌دهد. اصلاح تطبیق انجام‌شده با استفاده از دانش‌های نفی‌کننده به عنوان بعد سوم مجازی (اطلاعات غنی‌کننده) در فرآیند تصمیم‌گیری فیلتر ذره‌ای تأثیر می‌گذارد و نتیجه نهایی را اصلاح می‌کند. درست مانند دانشی که به طور قطع از عدم تعلق یک موجودیت به یک گروه صحبت می‌کند؛ بنابراین مسئله مهم و اصلی در اینجا یافتن چنین دانشی است. با علم به اینکه تنها دو بعد تصویر فضای مسئله ماست، یافتن چنین دانشی نیاز به بررسی دارد.

۲-۵. ساختار سامانه پیشنهادی

این ساختار در دو بخش، بررسی می‌شود:

سرعت شیء، اندازه شیء، رنگ شیء در فضای رنگی RGB. بافت شیء و میزان تفاضل با کادر محاطی و میزان کشیدگی سایه.

جدول (۱). لیست کلیه ویژگی‌هایی که می‌توان از اشیای متحرک صحنه‌های دوبعدی یک دنباله تصویر ویدئویی استخراج کرد

ویژگی‌های ظاهری شیء متحرک	ویژگی‌های حرکتی شیء متحرک
رنگ	جهت حرکت
شدت یا غلظت رنگ	سرعت حرکت
بافت	فاصله از کناره‌ها
شکل کلی	تغییرات اندازه شیء
تعداد اشکال تشکیل‌دهنده شیء	مدل حرکتی
نوع اشکال تشکیل‌دهنده شیء	تغییرات اندازه سایه
مساحت	
میزان تفاضل با کادر محاطی	
کشیدگی سایه	
پیوستگی سایه به جسم	
جهت سایه	
نسبت کشیدگی افقی/عمودی	

رنگ: تنها برای محیط‌های با نور ثابت و اشیایی که از هر طرف یک رنگ هستند، مناسب است.

بافت: برای اشیای با بافت ثابت مناسب است.

شکل کلی: در صورتی که انسداد و هم‌پوشانی‌ها زیاد نباشد، مناسب است.

نوع و تعداد اشکال: در صورتی که انسداد و هم‌پوشانی‌ها زیاد نباشد، مناسب است.

کشیدگی سایه: باتوجه به حرکت شیء و دور و نزدیک شدن به منابع نور مناسب نیست.

پیوستگی سایه به شیء: مناسب به نظر می‌رسد.

جهت سایه: باتوجه به حرکت شیء و دور و نزدیک شدن به منابع نور مناسب نیست.

جهت حرکت: به هیچ وجه مناسب نیست.

سرعت حرکت: به تنهایی مناسب نیست.

میزان تغییرات اندازه: به دلیل دخالت زمان و سرعت مناسب به نظر می‌رسد.

میزان تغییرات اندازه سایه: مناسب نیست.

میزان نسبت کشیدگی افقی/عمودی: برای اجسام صلب مناسب به نظر می‌رسد.

با وجود استدلال‌های بالا، انتخاب ۳ ویژگی برتر، همانند انتخاب ویژگی‌های مربوط به فیلتر ذره‌ای با آزمایش‌های متعدد انجام شد. نتیجه آزمایش‌ها با استدلال-های انجام گرفته هم‌خوانی داشت. نتیجه این آزمایش‌ها در نمودارهای شکل (۶) و (۷) و (۸) آورده شده است. نمودار شکل (۶) شش ویژگی قابل استفاده در فرآیند تخمین احتمال شرطی پسین فیلتر ذره‌ای و تأثیر هر یک در دقت ردیابی را نشان می‌دهد. این نرخ تأثیر در دقت با آزمایش فیلتر ذره‌ای معمول و حذف هر بار یک ویژگی از شش ویژگی و ثبت نتیجه به دست آمده است. شکل (۷)

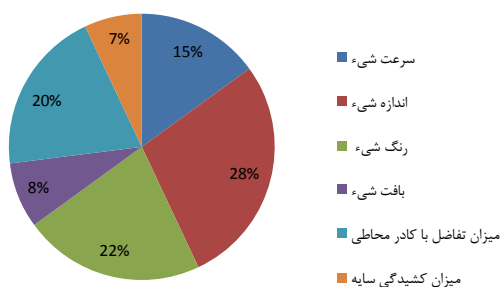
به منظور عملکرد بی‌درنگ با توجه به آزمایش‌های انجام گرفته، به‌طور متوسط استخراج سه ویژگی از شش ویژگی انتخابی مناسب است. برای ویژگی‌های نفی‌کننده نیاز به ویژگی‌هایی داریم که مقادیر آن‌ها تغییرات کم یا ثابتی برای یک شیء خاص داشته باشند تا عدم برقراری آن مقادیر، تعلق موجودیتی با آن ویژگی به شیء خاص را نفی کند. در ادامه، کلیه ویژگی‌های ممکن را معرفی کرده، در مورد امکان‌پذیری استفاده از آن‌ها به این منظور، بحث می‌کنیم.

درنهایت، ویژگی‌های نشان داده شده در جدول (۲) برای هر بخش انتخاب شده است. میزان نوآوری هر ویژگی یا مرجع معرفی آن، چگونگی استخراج آن و وزن هر ویژگی نیز در جدول (۲) نشان داده شده است.

نمودار مشابهی را برای ویژگی‌های قابل به کارگیری به عنوان اطلاعات غنی کننده به عنوان دانش افزوده و تأثیر هریک در نرخ دقت ردیابی را نشان می دهد. این میزان تأثیر با آزمایش مستقل هر ویژگی به عنوان دانش غنی کننده و مقایسه نتایج به دست آمده است. شکل (۸) نمودار مقایسه میزان دقت ردیابی در ترکیب‌های متفاوت سه تایی از هر دسته ویژگی را نشان می دهد. با توجه به این نتایج،

جدول (۲). ویژگی‌های مورد استفاده در روش پیشنهادی، مرجع هریک و وزن نرمال هر ویژگی در فیلتر ذره‌ای

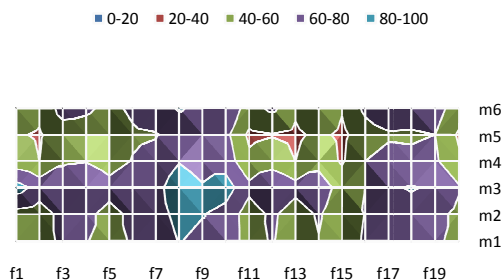
ویژگی‌های به کار رفته در فیلتر ذره‌ای			ویژگی‌های به کار رفته برای سؤالات نفی کننده		
ویژگی	مرجع	چگونگی استخراج	ویژگی	مرجع	چگونگی استخراج
شکل با وزن ۰.۸	[۲۸]	تطبیق با اشکال اولیه و یافتن بیشترین تطبیق	اندازه نسبی	[۱۹]	تفاضل مساحت مستطیل محاطی شیء و مساحت شیء
اندازه با وزن ۰.۶	[۳۴]	نسبت تعداد نقاط اشغالی توسط شیء به کل نقاط قاب	پیوستگی با سایه	-	نسبت تعداد نقاط مرزی شیء به تعداد نقاط مرزی با همسایگی سایه
رنگ با وزن ۰.۴	[۶]	هیستوگرام میانگین مؤلفه‌های رنگی در فضای RGB	نسبت کشیدگی افقی / عمودی	-	نسبت تفاضل بالاترین نقطه از پایین ترین نقطه به تفاضل چپ ترین و راست ترین نقاط



۴-۵. ساختار فیلتر ذره‌ای غنی شده پیشنهادی

مراحل زیر ساختار چارچوب ردیابی به کمک فیلتر ذره‌ای غنی شده را پس از استخراج ویژگی‌ها ایجاد می کند:

شکل (۶). ویژگی‌های مناسب برای استفاده در فرآیند تخمین احتمالات پسین در فیلتر ذره‌ای. تنها این شش ویژگی است که به علت قابلیت تخمین از روی تاریخچه مشاهدات، برای این امر مناسب تر به نظر می رسد. ویژگی‌ها از سرعت شیء با میزان تأثیر ۱۵٪ به صورت ساعتگرد در نمودار چیده شده است.



شکل (۸). مقایسه نرخ دقت ردیابی در ترکیبات سه تایی از هر دسته ویژگی. محور عمودی نمودار شش ترکیب سه تایی ممکن از شش ویژگی مربوط به فیلترزده ای عادی و محور افقی ترکیبات سه تایی مناسب از ۱۱ ویژگی مربوط به اطلاعات غنی کننده را نشان می دهند. تلاقی m3 و f8 بیشترین نرخ دقت را ارائه می دهد. ترکیب سه تایی m3 و f8 در جدول (۲) معرفی شده است.

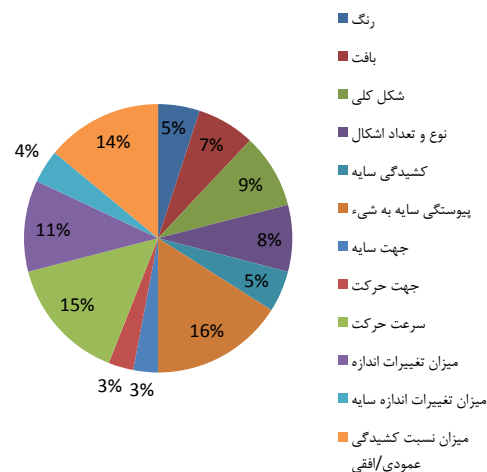
(ب) برای هر ذره در قاب جاری ویژگی های معرفی شده در جدول ۲ استخراج می شود.

For i=1 to n when n= num of (O_i^n) do (۱۰)
For j=1 to 3 do
Extract features from $O_{20}^{i,j}$

(ج) این ویژگی ها برای هر ذره به عنوان پارامترهای متغیر حالت هر ذره استفاده می شود.

For i=1 to n when n= num of (O_i^n) do (۱۱)
For j=1 to 3 do
 $X_{20}^{i,j} = (x, y, o, v, z)$

(د) احتمال حالت پسین هر ذره و مقادیر متغیرهای حالت هر ذره با استفاده از روش انتشار تابع چگالی احتمال و معادلات (۵) تا (۸) محاسبه می شود. درواقع این محاسبات حالت پسین شیء را در مورد ویژگی های در نظر گرفته شده به دست می دهد.



شکل (۷). ویژگی های کاندید برای به کارگیری به عنوان اطلاعات غنی کننده به عنوان دانش افزوده. ویژگی هایی در این بخش، میزان تأثیر بیشتری در دقت دارند که در طول دنباله تصاویر، برای شیء متحرک تقریباً ثابت باقی بمانند. ویژگی ها در نمودار از ویژگی رنگ با نرخ تأثیر ۵٪ به صورت ساعتگرد چیده شده اند.

(الف) ردیابی از قاب دهم از دنباله تصاویر و پس از گذشت ۲۰ قاب یعنی در حدود ۱ ثانیه پس از شروع دنباله تصاویر شروع می شود. در این قاب هر شیء متحرک i به سه بخش مساوی افقی یا عمودی بسته به چگونگی کشیدگی شیء تقسیم شده و به هر بخش یک شناسه ذره تعلق می گیرد.

If (i in $F_i = 20$) then (۹)
For i=10 do
For j=1 to n when n= num of (O_i^n) do
Divide O_i^j to $O_i^{j,1}, O_i^{j,2}, O_i^{j,3}$

برای ذرات منطبق شده با حد آستانه مناسب که سؤالات نفی کننده انطباق آن‌ها را تأیید می کنند، ذره جدید به عنوان مدل مشاهده در تاریخچه ذره قرار گرفته و طبق معادله (۶) وزن آن ذره اصلاح می شود.

$$\begin{aligned} & \text{If}(\text{compare}(o,v,z, \\ & \ddot{x}_{n+1}^{i,j}, \ddot{x}_{n+1}^{i,j}) \text{ verify the match the} \\ & \ddot{x}_{n+1}^{i,j} = x_{n+1}^{i,j} \text{ and} \\ & w_k^i = p(z_k | x_k = x_k^{(i)}) = L(z_k | x_k = x_k^i) \end{aligned} \quad (15)$$

چنانچه انطباق ذره‌ای پس از بخش ی نفی شد، مرحله و بدون ذره جاری تطبیق داده شده انجام می گیرد.

برای برخی از ذرات که با هیچ احتمال پسینی در نهایت منطبق نمی شوند، طبق معادله (۶) وزن ذرات کاهش پیدا می کند. چنانچه وزن هر ذره از حد آستانه در نظر گرفته شده کمتر شد، آن ذره از مجموعه کل ذرات حذف می شود. این ذرات مربوط به اجسامی است که در طول دنباله تصاویر از میدان دید خارج شده و یا ثابت شده اند. میزان حد آستانه چنانچه خیلی کم باشد، باعث کندی و چنانچه بزرگ باشد باعث کاهش کارایی در اثر گم کردن شیء با کوچک ترین سکون یا خروجی می شود.

ح) برای همه ذرات جدید احتمال پسین حالت آن‌ها محاسبه شده و به مجموعه احتمالات پسین موجود افزوده می شود.

ط) برای هر قاب بعدی از مرحله ۵ به بعد تکرار می شود.

الگوریتم کامل روش پیشنهادی در شکل (۹) نمایش داده شده است.

$$\text{For each } x_n^{i,j} : x_{n+1}^{i,j} = P(x_{n+1}^{i,j} | x_n^{i,j}, z_{n+1}^{i,j}) \quad (12)$$

ه) در قاب بعدی مراحل (الف) تا (ج) تکرار شده و برای اشیای متحرک مقادیر ویژگی های هر ذره محاسبه می شود. و طبق احتمالات پسین حالات ذرات قاب قبلی، میزان شباهت هر احتمال پسین با هر ذره طبق فاصله معروف بهاتاچاریا محاسبه می شود. ضریب شباهت به دست آمده عددی بین صفر و یک است و مقادیر بزرگ تر از ۰.۸ برای آن، ذره جدید را با ذره قبلی و در نتیجه شیء قبلی منطبق می کند.

$$\begin{aligned} & \text{For each } x_{n+1}^{i,j} \text{ and each } \ddot{x}_{n+1}^{i,j} \quad (13) \\ & \text{Calculate } \rho(x_{n+1}^{i,j}, \ddot{x}_{n+1}^{i,j}) = \\ & \sum_{i=1}^n \sqrt{x_{n+1}^{i,j} \ddot{x}_{n+1}^{i,j}}, \sum_{i=1}^n \ddot{x}_{n+1}^{i,j} = 1 \\ & \text{If } \rho > .08 \text{ then } \ddot{x}_{n+1}^{i,j} = x_{n+1}^{i,j} \end{aligned}$$

ایده آل آن است که تمامی ذرات یک شیء به ذرات یک شیء واحد منطبق شوند. در برخی موارد که اشیا دچار انسداد یا هم پوشانی شوند، ممکن است این اتفاق به درستی رخ ندهد.

ز) در این مرحله، سؤالات نفی کننده برای تأیید یا رد انطباق انجام شده وارد عمل می شوند؛ به عبارت دیگر اشیای ذراتی که بر هم منطبق شده اند از لحاظ اندازه، پیوستگی سایه و نسبت کشیدگی افقی/ عمودی مقایسه می شوند. برای این مقایسه نیز از ضریب شباهت بهاتاچاریا استفاده شده است.

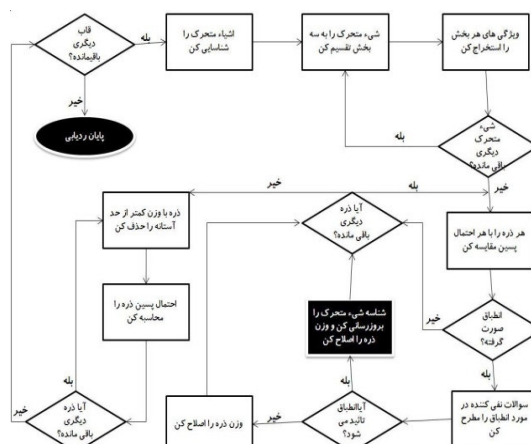
$$\begin{aligned} & \text{For each } \ddot{x}_{n+1}^{i,j} \text{ and each } \ddot{x}_{n+1}^{i,j} \quad (14) \\ & \text{Compare}(o,v,z, \ddot{x}_{n+1}^{i,j}, \ddot{x}_{n+1}^{i,j}) \end{aligned}$$

نتایج این آزمایش‌ها بر روی مجموعه تصاویر نام-برده‌شده، در جدول (۵) آمده است. از اطلاعات این جدول میزان برتری روش پیشنهادی در هر سه دسته از تصاویر ویدئویی کاملاً مشهود است. بیشترین افزایش دقت مربوط به صحنه‌های نیمه‌شلوغ است که در حدود ۱۲.۵٪ دقت ردیابی نسبت به بهترین دقت سه روش دیگر، بهبود حاصل شده است. بیشترین کاهش نرخ خطای ردیابی نیز مربوط به همین گروه صحنه‌ها و در حدود ۱۲.۶٪ بوده است.

در جدول (۵) همچنین پیچیدگی زمانی روش‌های مختلف بر حسب زمانی که برای پردازش یک قاب در هر روش صرف می‌شود، نشان داده شده است. تقریباً تمامی آن‌ها پیچیدگی‌های زمانی مشابهی در هر گروه داشته و به‌صورت تقریباً بی‌درنگ فرآیند ردیابی را انجام می‌دهند؛ زیرا بخش دانش‌افزوده در بعد اطلاعات غنی‌کننده و محاسبات آن در تعیین میزان قطعیت تخمین، فرآیندی با پیچیدگی خطی بوده و تأثیر چندانی در زمان ردیابی نخواهد داشت.

نمودارهای شکل (۱۰) و شکل (۱۱) نرخ دقت و خطا را برای مجموعه داده‌های PETS09 نمایش می‌دهد. روش پیشنهادی در تصاویر شلوغ که اغلب اشیاء، انسان هستند، نسبت به روش‌های موجود، کارایی بالاتری از خود نشان می‌دهد. علت این امر، مقاومت بیشتر روش در برابر تغییرات حالات تصادفی اجسام و گم‌نشدن شیء در انسدادها و هم‌پوشانی‌هاست.

نمونه‌ای از نتایج خروجی در برخی از قاب‌ها در شکل (۱۲) نشان داده شده است.



شکل (۹). فلوجارت فرآیند ردیابی در روش پیشنهادی. مستطیل مشکی رنگ جایی است که در آن با به روز رسانی شناسه شیء متحرک، فرآیند ردیابی برای هر قاب انجام می شود. چارچوب ردیابی فیلتر ذره ای ارائه شده توسط [۲۹] به صورت جعبه ابزار مبتنی بر استفاده از ویژگی های مورد استفاده در روش پیشنهادی

۶. نتایج آزمایش‌ها

برای آزمایش روش و مقایسه آن با روش‌های مشابه موجود، مجموعه داده‌های خود را در سه گروه مطابق جدول (۳) تقسیم‌بندی کرده‌ایم. همچنین با پیاده‌سازی سه روش دیگر، به مقایسه کارایی روش خود با روش‌های مشابه پرداخته‌ایم. این روش‌ها عبارت‌اند از:

- چارچوب فیلتر ذره‌ای معرفی شده در [۲۸] که از ویژگی‌های مکان و رنگ شیء متحرک استفاده می‌کند.

– چارچوب ردیابی فیلتر ذره‌ای ارائه‌شده در [۲۹] به‌صورت جعبه ابزار Matlab با استفاده از ویژگی‌های مورد استفاده در روش پیشنهادی.

– چارچوب فیلتر ذره‌ای معرفی شده در روش پیشنهادی بدون استفاده از اطلاعات غنی‌کننده.

۶-۱. مزایا و معایب روش پیشنهادی

از جمله مواردی که می‌توان مزیت روش پیشنهادی نسبت به روش‌های دیگر برشمرد عبارت‌اند از:

عملکرد تقریباً بلادرنگ نسبت به روش‌های دیگر فیلتر ذره‌ای که دارای تعداد زیادی ذره هستند.

مقاومت در برابر هم‌پوشانی اشیای متحرک به‌خصوص در مواردی که حالت بعدی آن‌ها در برخی ویژگی‌ها تصادفی است. مقاومت در برابر تغییرات شرایط محیطی.

عدم افت زیاد کارایی با افزایش تعداد اشیای متحرک در صحنه.

در کنار این مزایا هنوز نواقصی در روش پیشنهادی باقی مانده است که از جمله می‌توان به موارد زیر اشاره کرد:

وابستگی روش به حدود آستانه انتخابی که باعث می‌شود کارایی روش از یک دنباله تا دنباله تصویر دیگر متفاوت باشد.

امکان استفاده از ضرایب بهتر و یا حتی روش‌های دسته‌بندی در تعیین میزان شباهت ذرات که در روش پیشنهادی لحاظ نشده است.

استفاده از توابع احتمالاتی توسعه‌یافته به جای تابع معمول انتشار چگالی احتمال در فیلتر ذره‌ای.

انتخاب ذرات به‌صورت ثابت که ممکن است در برخی موارد با تعداد دیگری کارایی بهتری به‌دست آید.

به‌طور کلی می‌توان گفت که امکان بهبود روش پیشنهادی با همین ایده ولی استفاده از ابزارها و روش‌های دیگر در بخش‌های مختلف آن امکان‌پذیر است.

جدول (۴). معرفی دقیق معیارهای دقت و خطای ردیابی

معیار	توضیح
دقت $\eta = \frac{TP_p}{TP_p + FN_p}$	نشان‌دهنده توانایی سامانه ردیاب در نگاشت درست هر ذره به شیء متعلق به آن. به عبارت دیگر، نسبت ذرات درست تشخیص داده‌شده به کل ذرات درست
خطا $\xi = \frac{TP_f}{TP_f + FN_f}$	نشان‌دهنده خطای روش در نگاشت ذرات به هم، به عبارت دیگر، تعداد ذرات جدیدی که به اشتباه نگاشت نشده‌اند به علاوه تعداد ذراتی که به اشتباه نگاشت نشده‌اند به مجموع ذرات
TP_x : مقادیری از موجودیت x که به درستی موجودیت x تشخیص داده شده‌اند FN_x : مقادیری از دیگر موجودیت‌ها که به اشتباه موجودیت x تشخیص داده شده‌اند	

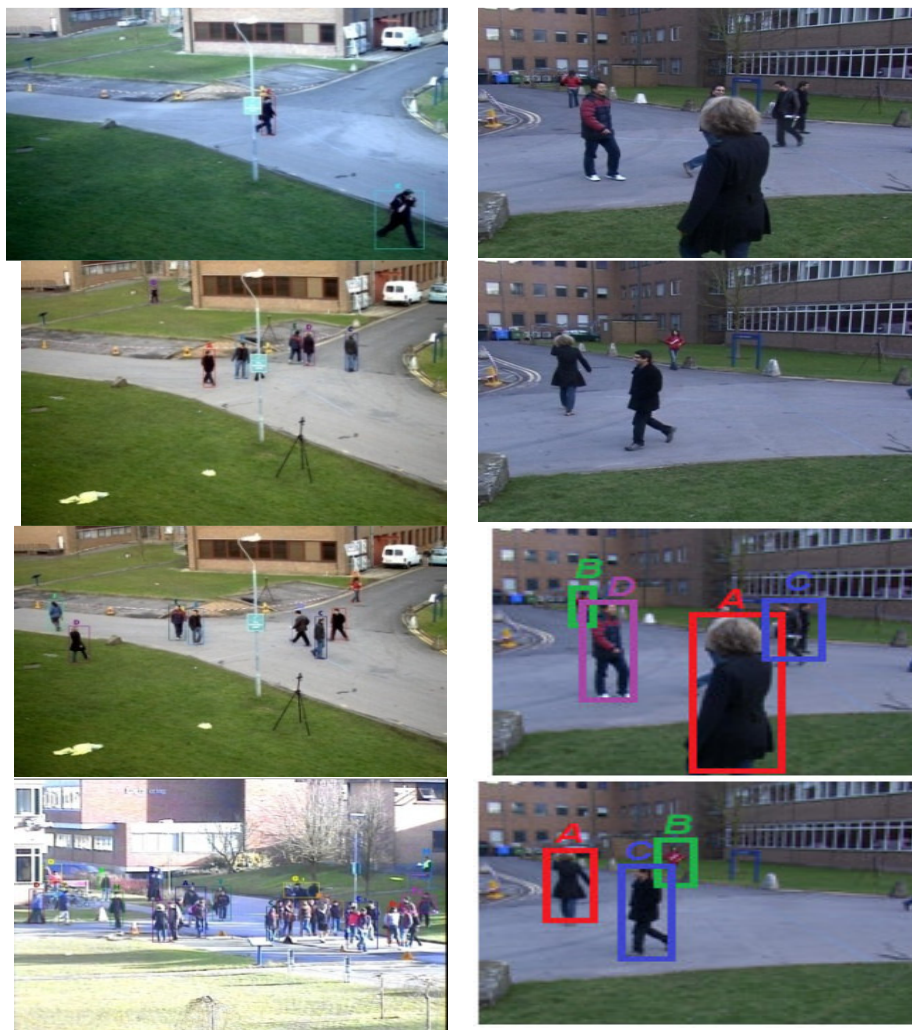
۷. جمع‌بندی

در این مقاله روش پیشنهادی خود را برای ردیابی اشیای متحرک در صحنه‌های نیمه‌شلوغ با یک دوربین ثابت معرفی کرده‌ایم. این روش، مبتنی بر چارچوب احتمالاتی فیلتر ذره‌ای است که با ایده به‌کارگیری اطلاعاتی اضافی به عنوان اطلاعات غنی‌کننده برای درک بهتر محیط عمل می‌کند. از آنجا که هیچ‌گونه اطلاعات دیگری جز دو بعد

تصویری در دنباله قاب‌های تصویر در دسترس نیست، ناچار به استفاده از برخی داده‌های استخراجی از تصویر به عنوان اطلاعات غنی‌کننده بوده‌ایم. در این پژوهش، از روش منطقی سؤالات نفی‌کننده به عنوان دانش اضافی بهره برده‌ایم. مشکل اصلی در به‌کارگیری این ایده، یافتن سؤالات نفی‌کننده مناسب است که در این قسمت از پژوهش از میان ویژگی‌های قبلی و برخی ویژگی‌های

مناسب‌تر استفاده کرد، می‌توان از ایده مطرح شده انتظار کارایی بیشتر را نیز داشت.

جدید معرفی شده، با انجام آزمایش انتخاب شده‌اند. مزیت اصلی این روش، در مقاومت فرآیند ردیابی در برابر انسداد و هم‌پوشانی اشیای مشابه، حرکت تصادفی اشیای متحرک و تغییرات حالت‌های اشیای متحرک است. چنانچه بتوان از ویژگی‌های بهتر و یا ضرایب شباهت



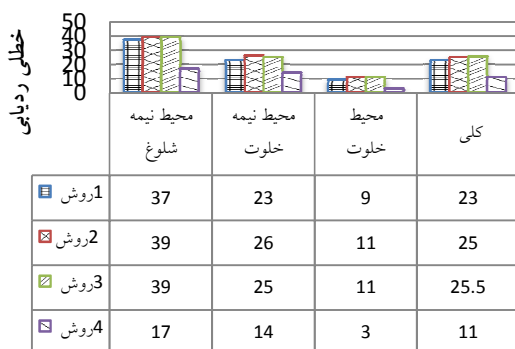
شکل (۱۲). نمونه‌هایی از نتایج خروجی روش پیشنهادی در ستون سمت چپ، به ترتیب از سمت بالا به پایین، محیط خلوت، محیط نیمه-خلوت، محیط نیمه‌شلوغ و محیط شلوغ (که به‌وضوح کارایی روش در آن افت کرده است). در ستون سمت راست، به ترتیب از بالا، قاب شماره ۴۵۵، قاب شماره ۶۹۰، اشیای ردیابی شده در قاب ۴۵۵ام با شناسه آن‌ها و اشیای شناخته‌شده در قاب ۶۹۰ام با شناسه آن‌ها با استفاده از روش پیشنهادی نشان داده شده است. در قاب ۴۵۵ام دو شیء کنار هم، به اشتباه یک شیء در نظر گرفته شده است؛ اما در قاب ۶۹۰ام پس از جدا شدن اشیاء، فرآیند ردیابی به‌درستی ادامه پیدا کرده است.

جدول (۳). مشخصات محیط‌های مختلف مورد استفاده در آزمایش فرآیندهای ردیابی

	بیرونی و درونی	بیرونی و درونی	بیرونی و درونی
	۲۳۰۰	۱۴۵۰	۹۰۰
	320*240	320*240	320*240
	انسان	انسان+وسیله نقلیه	انسان
	متوسط	متنوع	متوسط
	خیلی بالا	بالا	پایین
	خیلی بالا	بالا	متوسط
	کم	خیلی کم	خیلی کم

جدول (۵). نتایج آزمایش‌های مربوط به محاسبه نرخ خطا و دقت ردیابی و همچنین زمان پردازش هر قاب در روش‌های مختلف روش‌های اول تا سوم روش‌های بلادرنگ بوده و با پیچیدگی زمانی $O(n)$ ، هر قاب را در $n/1$ یا $n/3$ ثانیه پردازش می‌کند؛ اما روش پیشنهادی در محیط‌های نیمه خلوت و نیمه شلوغ کمی کندتر از روش‌های دیگر عمل می‌کند و درواقع هر قاب را در $n/2$ و $n/3$ ثانیه پردازش می‌کند. پیچیدگی زمانی این روش نیز در حالت کلی $O(n)$ است؛ ولی در حالت دقیق از مرتبه $\theta(n+2)$ و $\theta(n+1/2)$ است. درواقع روش پیشنهادی، در بدترین حالت از هر ۳ قاب تصویر، ۱ قاب را به صورت بلادرنگ پردازش می‌کند.

	T:fps			T:fps			T:fps		
	%۶۸	%۳۸		%۷۹	%۱۹		%۹۱	%۷	
	%۶۵.۵	%۳۴.۴		%۷۸	%۱۹		%۸۷	%۱۱	
	%۶۳	%۴۱		%۷۵	%۲۳		%۸۳	%۱۵	
	%۸۰.۵	%۲۳	3	%۸۵	%۱۶		%۹۶	%۳.۵	



شکل (۱۱). نرخ خطای ردیابی در روش‌های مختلف و محیط‌های مختلف برای مجموعه PETS09. آخرین ستون این نرخ را به صورت کلی نشان می‌دهد.



شکل (۱۰). نرخ دقت ردیابی در روش‌های مختلف و محیط‌های مختلف برای مجموعه PETS09. آخرین ستون این نرخ را به صورت کلی نشان می‌دهد.

۸. منابع

- MachineIntelligence*, vol. 32, no. 7, pp. 1239–1258, 2010..
- [12] Isard M., Blake A., "CONDENSATION-Conditional density propagation for visual tracking", *International Journal of Computer Vision*, 29:5-38, 1998.
- [13] Brasnett, P., et al., "Sequential Monte Carlo tracking by fusing multiple cues in video sequences". *Image Vision Comput.*, 2007. 25: p. 1217-1227.
- [14] McCane B., Galvin B., Novis K., "Algorithmic fusion for more robust feature tracking", *Intl. Journal Computer Vision*, 49:79-89, 2002.
- [15] Pérez, P., J. Vermaak, and A. Blake. "Data Fusion for Visual Tracking with Particles". in *Proc. IEEE*. 2004.
- [16] Mohedano, R.[Raul], del-Blanco, C.R.[Carlos R.], Jaureguizar, F.[Fernando], Salgado, L.[Luis], Garcia, N.[Narciso], "Robust 3D people tracking and positioning system in a semi-overlapped multi-camera environment", *ICIP08*(2656-2659).
- [17] Guo-liang Wang, De-qun Liang, Yan-chun Wang, Zhao-hua Hu, "Algorithm for Tracking of Fast Motion Objects with Adaptive Mean Shift," *snpd*, vol. 1, pp.359-363, Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD 2007)
- [18] Mohand Said Allili, Djemel Ziou, "Object Contour Tracking in Videos by Matching Finite Mixture Models," *avss*, pp.35, 2006 IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS'06), 2006
- [19] M. Asadi, A. Dore, A. Beoldo, C.S. Regazzoni, "Tracking by using dynamic shape model learning in the presence of occlusion," *avss*, pp.230-235, 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, 2007
- [20] Johansson, B.[Bjorn], Wiklund, J.[Johan], Forssen, P.E.[Per-Erik], Granlund, G.[Gosta], "Combining shadow detection and simulation for estimation of vehicle size and
- [1] A. Yilmaz, O. Javed, and M. Shah, "Object Tracking: A Survey" *ACMComput. Surv.*, vol. 38, no. 4, p. 13, 2006
- [2] GAVRILA, D. M. 1999. "The visual analysis of human movement: A survey." *Comput. Vision Image Under-*
- [3] Medeiros, H.[Henry], Park, J.[Johnny], Kak, A. C.[Avinash C.], "Distributed Object Tracking Using a Cluster-Based Kalman Filter in Wireless Camera Networks", *SelectedSP(2)*, No. 4, August, 2008, pp. 448-463
- [۴] ح. مرآتی, "ردیابی اشیای متحرک با استفاده از بهبود ردیابی با ماسک باینری" پایان‌نامه کارشناسی ارشد دانشگاه صنعتی شریف, ۱۳۸۵.
- [5] Gustafsson, F., Gunnarsson, F. and Bergman, N., "Particle filters for positioning, navigation, and tracking". *IEEE Trans. Signal Process.* v50 i2. 425-437.
- [6] Matej Kristan, "Tracking people in video data using probabilistic models" Ph. D Thesis University of Ljubljana Faculty of Electrical Engineering. April 2008
- [7] H. Chen, K. Zhang, and X. R. Li, "Optimal data compression for multisensory target tracking with communication constraints," in *Proc. 43rd Conf. Dec. Control, Atlantis, Bahamas, Dec. 2004*, pp. 2650–2655.
- [8] Huasheng et al., "Face tracking using skin-color model and gabor wavelet", *IEEE 0-7803-7488-6/02*, pp. 837-840.
- [9] Heikkila, M. and M.P.inen, "A Texture-Based Method for Modeling the Background and Detecting Moving Objects". *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2006. 28(4): p. 657-662.
- [10] Allili, M.S.[Mohand Saïd], Ziou, D.[Djemel], "Active contours for video object tracking using region, boundary and shape information", *SIViP(1)*, No. 2, June 2007, pp. 101-117.
- [11] D. Ger ´onimo, A. M. L´opez, A. D. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," *IEEE Transactions on Pattern Analysis and*

- players using the boosted particle filter", IVC(27), No. 1-2, January 2009, pp. 189-205
- [29] Hue, C.[Carine], Le Cadre, J.P.[Jean-Pierre], Pérez, P.[Patrick], "Tracking Multiple Objects with Particle Filtering", INRIARR-4033, October 2000.
- [30] Martin Godec, "Robust Object Tracking using Semi-Supervised Online Boosting", Master Thesis, Graz University of Technology, Institute for Computer Graphics and Vision October 2008
- [۳۱] مجید کمیلی، "ترکیب مناسب ویژگی‌ها برای ردیابی عابر بر مبنای فیلتر ذره در تصاویر مرئی" پایان‌نامه کارشناسی ارشد دانشگاه تربیت مدرس، تابستان ۱۳۸۷.
- [32] Khansari, M.; Rabiee, H.R.; Asadi, M.; Ghanbari, M. "Object tracking in crowded video scenes based on the undecimated wavelet features and texture analysis". EURASIP J. Adv. Signal Process. 2008, Article ID 243534, 18 p. (2008). Hindawi Publishing Corporation, New York, NY
- [33] Jaward Mihaylova Canagarajah and M. Jaward and L. Mihaylova and N. Canagarajah and D. Bull, "Multiple Object Tracking Using Particle Filters" In IEEE CVPR, Madison, Wisconsin, June 2003
- [34] H. Moon, R. Chellappa, A. Rosenfeld, "3D Object Tracking Using Shape-Encoded Particle Propagation," iccv, vol. 2, pp.307, Eighth International Conference on Computer Vision (ICCV'01) - Volume 2, 2001
- [35] Sohail Nadimi, Bir Bhanu, "Moving Shadow Detection Using a Physics-Based Approach," icpr, vol. 2, pp.20701, 16th International Conference on Pattern Recognition (ICPR'02) - Volume 2, 2002
- [36] Bar-Shalom, Y., Kirubarajan, T., "Tracking with Classification-Aided Multiframe Data Association," Proceedings of SPIE Conference on Signal and Data Association of Small Targets 2003, vol. 5204.
- [37] <http://www.cvg.rdg.ac.uk/PETS2006/data.html>
<http://www.cvg.rdg.ac.uk/WINTERPETS09/a.html>
- position", PRL(30), No. 8, 1 June 2009, pp. 751-759.
- [21] Mahsa Maghami, Reza A. Zoroofi, Babak N. Araabi, Mohsen Shiva, Ehsan Vahedi, "Kalman Filter Tracking for Facial Expression Recognition using Noticeable Feature Selection," International Conference on Intelligent and Advanced Systems (ICIAS), Malasia, Nov. 2007.
- [22] K. Quast and A. Kaup, "Shape adaptive mean shift objecttracking using gaussian mixture models," in *Proceedings of the IEEE 11th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '10)*, Desenzano, Italy, 2010
- [23] K. Quast and A. Kaup, "Real-time moving object detectionin video sequences using spatio-temporal adaptiveGaussianmixture models," in *Proceedings of theInternational Conferenceon Computer Vision Theory and Applications (VISAPP '10)*, Angers, France, 2010
- [24] Maggio, E., F. Smerladi, and A. Cavallaro, "Adaptive multifeature tracking in a particle filtering framework". IEEE Transaction on Circuits and Systems for Video Technology, 2007. 17(10): p.1348-1359.
- [25] Kragic, D. and H.I. Christensen, Cue Integration for Visual Servoing. IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION, 2001. 17(1): p. 18-27.
- [26] Stanley T. Birchfield, Sriram Rangarajan, "SpatioGrams versus Histograms for Region-Based Tracking," cvpr, vol. 2, pp.1158-1163, 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2, 2005
- [27] J Ren, J. Orwell, G.A. Jones, M. Xu, "Tracking the Soccer Ball using Multiple Fixed Cameras" in 'Computer Vision and Image Understanding', Special Issue on Computer Vision Based Analysis in Sport Environments (2008)
- [28] Lu, W.L.[Wei-Lwun], Okuma, K.[Kenji], Little, J.J.[James J.], "Tracking and recognizing actions of multiple hockey