شبکه عصبی مبتنی بر تابع مبنای شعاعی برای کشف حرکات در محیط های پویا

خلاصه: کشف حرکات، فرایند بخش بندی اشیا متحرک در جریان های ویدئویی،اولین فرایند ضروری است و نقش مهمی در سیستم های نظارتی دارد. پویا بودن معمولا با شرایط درونی و بیرونی روبروست و شامل اشیائی چون پیچ وتاب خوردن درختان، فوران چشمه ها، موج دار شدن آب ، پرده های متحرک و غیره روبروست. با این حال، کشف دقیق و کامل حرکات در علم دینامیک اغلب یک وظیفه چالش برانگیز است. این مقاله یک رویکرد نوین کشف حرکات را بر اساس شبکه های عصبی، با تابع انحراف شعاعی ارائه می دهد که به دقت اشیا متحرک را نه تنها در علم دینامیک بلکه در علم استاتیک نیز کشف می کند. روش ارائه شده شامل دوماژول مهم است : ماژول ایجاد چند پس زمینه و ماژول بررسی اشیا متحرک. مدل ایجاد چند پس زمینه به طور کارامدی موجب ایجاد مدل احتمالی انعطاف پذیر می شود که از طریق فرایند یادگیری غیر نظارتی می تواند در پس زمینه های پویا یا ایستا کار کند. سپس ماژول بررسی شی متحرک درک کامل و دقیقی از بررسی شی متحرک ارائه می دهد، از طریق پردازش های بلوکی که احتمال دارد دربردارنده شی متحرک باشد. این امر از طریق دو رویه انجام می شود: رویه هشدار بلوکی و رویه استخراج شی. نتایج بررسی مدل ما به همراه مقایسات کمی و کیفی از طریق روشهای جدید، و بر اساس محدوده گسترده ای از توالی های ویدیویی عصبی ارائه می شود. نتایج کلی نشان می دهد که روش ارائه شده به طور نسبی بهتر از روشهای موجود عمل می کند با شباهت و نرخ های دقت F1 که به ترتیب 69.37 درصد و 65.5 درصد هستند.

**مقدمه :**

سیستم های نظارتی ویدیویی نیازمند تسهیل محدوده گسترده ای از برنامه های کاربردی در کامیپوتر هستند از جمله درک فعالیت انسانی، پایش ترافیک و تحلیل، محافظت از گونه های در حال انقراض و غیره. وظایف بسیاری در سیستم های نظارتی ویدیویی وجود دارد. این موارد عبارتند از (ولی محدود به این موارد نمیشوند) بررسی حرکت، طبقه بندی شی، ردیابی، شناخت رفتار و تحلیل فعالیت . اکتشاف حرکات، بخش بندی اشیا متحرک در جریان های ویدیویی است که اولین گام مرتبط است و نقشی مهم در سیستمهای نظارتی ویدیویی دارد.

رویکردهای مختلفی ارائه شده است تا کشف حرکات کامل و دقیقی وجود داشته باشد. سه طبقه بندی از رویکردهای کشف حرکات سنتی عبارتند از جریان اپتیکال، تفاوت موقتی و کاستن پس زمینه. جریان اپتیکال می تواند با بهبود بررسی بدست آید که در نتیجه پرتوافکنی حرکت روی یک صفحه تصویر با تقریب مناسب است. با این حال، جریان اپتیکال نسبت به نویز ها حساس است ودر برنامه های واقعی از نظر محاسباتی مقرون به صرفه نیست. تفاوت موقتی بررسی اشیا متحرک با محاسبه تفاوت میان فریم های متوالی است و می تواند با تغییرات محیطی به طور کارامدی همراه شود. با این وجود، شکل های بدست آمده از اشیا متحرک معمولا ناقص هستند بخصوص زمانی که اشیا متحرک در یک صحنه ساکن هستند و یا حرکت آرامی را از خود نشان می دهند. کاستن پس زمینه روش بسیار مشهوری برای کشف حرکات است. این امر اشیا متحرک را با کاستن تصویر فعلی ازمدل پس زمینه مرجع تصویر قبلی بررسی می کند. این روش اکتشاف بدلیل توانایی در استخراج اشیا متحرک و در عین حال پیچیدگی نسبی زمانی، بسیار مورد توجه قرار گرفته است. با این حال، کشف حرکات از طریق کاستن پس زمینه، بدلیل ایجاد یک مدل پس زمینه جانبی، منجر به نتایج بررسی ناقص از اشیا متحرک می شود. این امر می تواند به طور خاص حتی بدتر هم شود، آنهم زمانی که شی متحرک در صحنه های پویا بررسی می شود.

سالهاست که استفاده از مدل ترکیب گوسین به عنوان پس زمینه پویای مدل برای بررسی اشیا متحرک، پیشنهاد می شود. رویکردی است که بسیار مورد استفاده قرار می گیرد تا بتوان کشف حرکات را در صحنه های پویا کاربردی کرد. این رویکرد، هر پیکسل را به طور مستقل از ترکیب گوسین مدلسازی می کند و توسط یک تقریب آنلاین به روز می شود. روشهای مختلفی ارائه شده اند تا GMM قدیمی را بهبود بخشند. بسیاری از این روشها GMM را در کنار سایر رویکرد های شناخته شده ای چون جریان اپتیکال و تفاوت موقتی قرار می دهند و یا تکنیک های یادگیری و به روز رسانی پارامترها را بهبود می بخشند. اگرچه این نتایج نشان دهنده بهبود است، پیچیدگی محاسباتی آنان همچنان بالاست.

در 36 ، روش کاستن پس زمینه جدیدی ارائه شده است که مشکل پس زمینه پویا را حل میکند و از تکنیک های ایجاد پس زمینه پویا واقعی استفاده می کند. این تکنیک که براساس فیلتر میانه موقتی با میانگین فیلتری است که به طور نمایی وزن دهی شده است و می تواند اشیا متحرک را به طور خودکار با پیچیدگی پایین محاسباتی بررسی کند.

روش کاستن پس زمینه خودسامانده SOBS ، از معماری نقشه شبکه عصبی خودسازمانده استفاده می کند تا یک مدل پس زمینه برای کشف اشیا متحرک در محیط های پیچیده بسازد. یک نقشه عصبی خودسازمانده شامل بردارهای وزنی 3در 3 است که برای ایجاد مدل پس زمینه برای هر پیکسل رنگی استفاده می شود. بااین حال، این امر نیازمند حافظه قابل توجه و الزامات زمانی محاسباتی بسیاری است.

یک مدل پس زمینه بسیار فشرده از طریق کاربرد تکنیک کمی کردن/خوشه بندی بدست می آید که از روش کاستن پس زمینه کتابچه کد استفاده می کند که توالی آموزشی از دوره های زمانی بلند مدتی است که مشاهدات در آن صورت می گیرد.

برخلاف روشهایی که دربالا بیان شد، استخراج کننده پس زمینه دیداری، یک مدل پس زمینه را ایجاد می کند و یک مجموعه از ارزشهای پیکسل را آزمون می کند که از گذشته و یا همسایگی مجاور گرفته می شود. اشیا متحرک با تعیین تفاوت های میان مدل پس زمینه و پسیل های ورودی فعلی بررسی می شوند.

در مقایسه با روشهای ذکر شده قبلی، این مقاله نشان دهنده رویکرد کشف حرکات بر اساس تابع مبنای شعاعی و شبکه های عصبی مصنوعی است تا بتوان اشیا متحرک را در صحنه های پویا بخش بندی کرد. این روش می تواند به طور کارامدی تغییرات محیطی را پذیرفته و بررسی دقیق و کاملی از صحنه های پویا و ایستا داشته باشد. اساسا، شبکه عصبی RBF دارای توانایی مسیر دهی غیر خطی قوی و انعطاف سیناپسی لوکال از نورون هاست که ساختار شبکه ای ضعیفی دارد. این امر امکان کاربرد کشف حرکات را در هر محیط پویا وایستایی می دهد

این مقاله به شکل زیر سازماندهی می شود. بخش 2 نشان دهنده پیمایشی از چندین مطالعه منتشر شده جدید مرتبط است که در مقایسات ما استفاده می شود. روش کشف حرکات ما در بخش 3 توضیح داده می شود. نتایج عملی روش ما سپس با سایر روشهای جدید در بخش 5 مقایسه می شود. در نهایت بخش 6 دربردارنده نتیجه گیری ماست.

مطالعات مرتبط

در این بخش ما پنج روش جدید را روی کشف حرکات در محیط های پویا بررسی می کنیم: روش ایجاد پس زمینه پویای واقعی RFBG ، GMM ، SOBS ، و روش کاستن CB و روش استخراج کننده پس زمینه دیداری.

ایجاد پس زمینه پویای واقعی

RDBG از دو بیتمپ  استفاده می کند تا مدل پس زمینه انطباقی اش را ایجاد کند. بیت مپ  در بردارنده مدل پس زمینه انطباقی است، در حالیکه  ، آخرین فریم دوربین است. هر پیکسل در  تایمر زمانی بلند مدت  را دارد و تایمر زمانی کوتاه مدت  . تایمر  برای محاسبه تعداد فریم ها استفاده می شود که در آن پیکسل  ارزشهای مشابهی دارد. وقتی تفاوت میان  و ارزش پیکسل روودی  در تلورانس از قبل تعریف شده  قرار گیرد، تایمر بلند مدت  افزایش می یابد و  با ارزش پیکسل ورودی  جایگزین می شود.

تایمر  برای شمارش فریم ها استفاده می شود که درآن ارزش پیکسل ورودی  از  متفاوت است. وقتی تفاوت میان  و ارزش پیکسل ورودی  از تلورانس از قبل تعریف شده بالاتر باشد، ارزش پیکسل  با ارزش پیکسل ورودی  جایگزین می شود. به طور مشابه اگر ارزش پیکسل ورودی  متفاوت از  باشد،  به صفر باز می گردد. در غیر اینصورت  افزایش خواهد یافت.

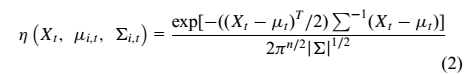
بنابراین اگر یک پیکسل توسط یک شی جدید پوشش داده شود،  باقی مانده و  افزایش می یابد. وقتی  بیشتر از  باشد، ارزش پیکسل ورودی  فرض می شود که بخشی از پس زمینه است. دراین مورد، ارزش پیکسل  و  با ارزش پیکسل ورودی  جایگزین شده و  به صفر باز می گردد.

مدل ترکیبی گوسین

استفاده از روش GMM شامل مدلسازی هر پیکسل به طور جداگانه می شود که ترکیبی از K گوسین را دارد و مدل پس زمینه احتمالی را حفظ می کند. تابع تراکم احتمال برای ارزش پیکسل فعلی به شکل زیر است:



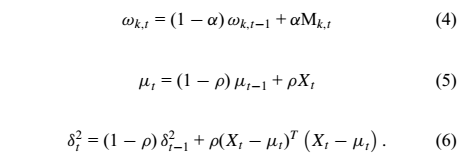
در جاییکه  ، هر ارزش شدت پیکسل روودی از t امین فریم تصویر است و برآر=ورد وزن ها برای توزیع گوسین متناظر است  که می تواند به شکل زیر بیان شود:



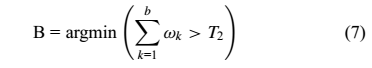
در جاییکه  و  ارزش های میانگین و ماتریس کوواریانس i امین گوسین در مدل ترکیبی هستند. ماتریس کوواریانس  به شکل زیر فرض می شود:



به یاد داشته باشید که هر پیکسل در برابر توزیع گوسین k موجود ارزیابی می شود. اگر ارزش پیکسل در 2.5 انحراف معیار یک توزیع باشد، وضعیت هماهنگ فرض می شود. پارامترهای انطباق از اولین مدل هماهنگ گوسین می تواند به شکل زیر به روز شود:



پارامتر  ، 1 است اگر پیکسل هماهنگ با مدل ها باشد. در غیر اینصورت صفر است. مدل های پس زمینه تطبیقی سس از طریق ارزش  از هر گوسین بدست می آیند و اولین توزیع B به شکل زیر تعیین می شود:



در جاییکه  ، نسبت حداقل داده هایی است که می توانند به عنوان پس زمینه طبقه بندی شوند.

کاستن پس زمینه خودسازمانده

روش SOBS شامل دو مرحله اصلی می شود. مرحله اول، مدل پس زمینه اولیه ای که بوسیله مسیر دهی به هر پیکسل تصویر اصلی ایجاد می شود که یک ماتریس ساختار مسیر عصبی 3 در 3 است. سپس، بهترین گزینه پس زمینه هماهنگ در ماتریس 3 در 2 هر پیکسل ورودی توسط آستانه ای ثابت تعیین می شود:



در جاییکه  پیکسل ورودی است ،  ، i امین کاندیدا در ماتریس 3 در 3 است و  بهترین هماهنگی است. اگر هیچ دو انطباقی با پیکسل ورودی  پیدا نشود،  بخش مشخصی از شی متحرک نامیده می شود. در غیر اینصورت  به عنوان یک پیکسل پس زمینه در نظر گرفته می شود. اگر  بهترین انطباق باشد در جایگاه  در مدل پس زمینه، مدل پس زمینه به شکل زیر به روز می شود:



د رجاییکه A مدل پس زمینه مسیر عصبی است و  نرخ یادگیری می باشد.

کسر پس زمینه Codebook

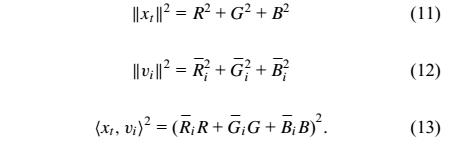
روش CB ، پس زمینه را با استفاده از تکنیک کمی سازی/خوشه بندی مدلسازی می کند که بر اساس مشاهداتی است که در طول زمان انجام می گیرد. CODEBOOK شامل یک یا چند کدواژه میباشد که که برای هر پیکسل ایجاد می شود. به شرح زیر است:



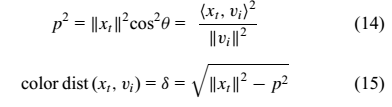
در جاییکه C شامل کتاب کد از L واژه کد است،  ، i امین واژه کد است که شامل بردار RBG  و  چندتایی است که می تواند به شکل  و  نشان داده شود.

به یاد داشته باشید که  به ترتیب حداقل و حداکثر ارزشهای شدت شفافیت از واژه های کد هستند. F نشان دهنده فراوانی است که در آن کدواژه ا شکل می گیرند.  بلند ترین فاصله زمانی دوره آموزشی است که در آن کدواژه ها رخ نمی دهند. p و q برای ثبت اولین و خرین زمان دسترسی به ترتیب مورد استفاده قرار میگیرند.

وقتی پیکسل ورودی  و RGB وجود دارد و  بردار RGB از کدواژه های  است، می توان به شکل زیر آن را بیان کرد:



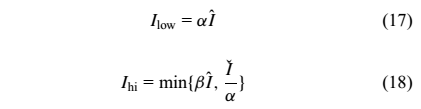
بنابراین توزیع رنگ به شکل زیر خواهد بود



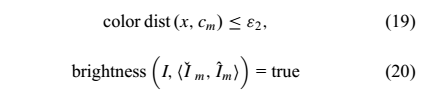
در جاییکه ، بردار RGB از کدواژه  است و  توزیع رنگ میان  و  است. تابع شفافیت منطقی به شکل زیر تعریف می شود:



بعلاوه محدوده  برای هر کدواژه به شکل زیر تعریف می شود:



درجاییکه  ، پارامترهای از قبل تعریف شده هستند.  محدوده ای از 0.4 و 0.7 را دارد و ، محدوده ای از 1.1 تا 1.5 را دارد. در نهایت نتایج بررسی می تواند از طریق دو شرط زیر بدست آید:



در جاییکه  آستانه بررسی است و  کدواژه پس زمینه است. اگر پیکسل ورودی در هر دو شرایط متناسب باشد، به عنوان پیکسل پس زمینه در نظر گرفته می شود؛ از سوی دیگر بخشی از شی متحرک است. بر اساس تجربیات انجام شده در مطالعات قبلی، میانگین تعداد کدواژه ها برای هر پیکسل برای رسیدن به پس زمینه 6.5 است.

استخراج کننده پس زمینه دیداری

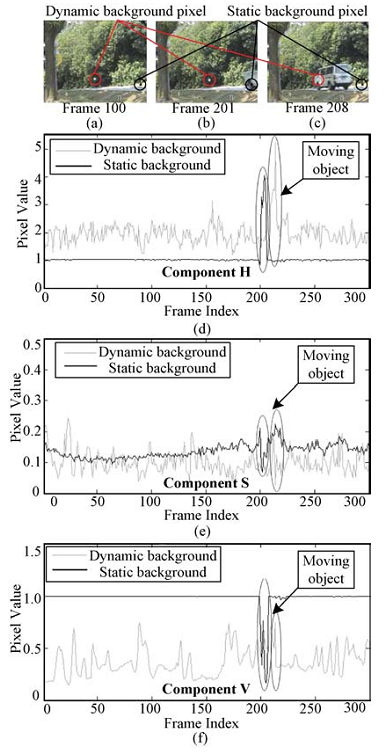
رویکرد VIBE اشیا متحرک را با محاسبه تفاوت میان مدل پس زمینه  و پیکسل ورودی  بررسی می کند. به طور اولیه، مدل پس زمینه از اولین فریم بدست می آید. T امین نمونه پس زمینه  به طور تصادفی با استفاده از N پیکسل مجاور در 8 همسایگی مرتبط مکانی  بدست می آیدو

دوم آنکه، انطباق مناسب زمانی رخ می دهد که فاصله اقلیدسی میان  کمتر از آستانه تعریف شده R باشد. اگر تعداد رویدادهای انطباق مناسب بزرگتر و یا مساوی هر آستانه مشخص  باشد، پیکسل های فعلی  به عنوان پس زمینه دسته بندی می شوند. از سوی دیگر  به عنوان یک پیکسل جلوی زمینه در نظر گرفته می شود.

در نهایت اگر  به عنوان پیکسل پس زمینه تعیین شود، دو نمونه پس زمینه که به طور تصادفی انتخاب شده اند یکی از مکان  و دیگری در مکان 8 نقطه ای مجاور مرتبط، با  جایگزین می شوند.

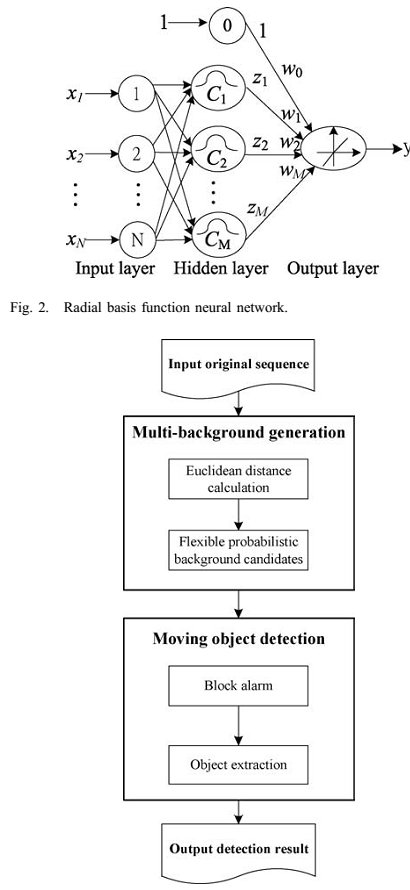
3. رویکرد ارائه شده مبتنی بر RBF

در کل، اکثر روشهای موجود می توانند به خوبی در شرایط ایستا عمل کنند. با این حال، کشف حرکات دقیق در محیط های پویا همچون پیچ و تاب خوردن درختان، آب های خروشان و غیره هنوز هم کار دشواری است. اولین دلیل برای این امر دشواری درتمایز میان اشیا متحرک و پس زمینه متحرک است که موجب تلاطم شدیددر پیکسل های پس زمینه و جلوی زمینه می شود



برای مثال همانطور که شکل 1 نشان می دهد، دو پیکسل توسط نقاط سیاه و سفید مشخص شده اند که برای کشیدن واریانس های شدت در عناصر hue، اشباع (s) ارزش (v) برای فریم های 500 از توالی نمونه هستند. نقاط سفید متعلق به درختان در حال حرکت است که به عنوان پیکسل پس زمینه پویا در شکل 1 a –c نشان داده شده اند. به طور معکوس، نقطه سیاه پیکسل های پس زمینه ایستا هستند. در شکل 1b , c ، چرخها از وسط پیکسل های پس زمینه پویا و پیکسل های پس زمینه ایستا به ترتیب می گذرند. شکل 1 d-f نشان دهنده نقاط متغیرهای شدت از دو پیکسل نمونه است که با نقاط سیاه و سفید برای عناصر H,S,V نشان داده شده اند. این امر نشان می دهد که سیگنال های پیکسل پس زمینه پویا ثابت هستند که امکان استخراج ساده شی متحرک را بدانها می دهد، زمانی که چرخه ها از مابین این پیکسل ها عبور میکند. با این حال، بدلیل نوسان های سیگنال ها ، دشوار است تا تمایزی میان سیگنال شی متحرک و پس زمینه متحرک پیدا کرد

در این بخش، ما یک رویکرد کشف حرکات جدید را بر اساس شبکه های عصبی مصنوعی RBF ارائه می دهیم. این فرایند در بردارنده اشیا متحرک از محیط های پویا و پس زمینه های ایستا است تا بتوان به طور کارامدی از قضاوت نادرست در خصوص پس زمینه های اشیا متحرک اطلاع پیدا کرد. شبکه عصبی RBF که در شکل 2 نشان داده شده است شامل یک لایه ورودی ، یک لایه نهان و یک لایه خروجی است. همچنین مزیت های خاصی وجود دارد که شامل تنظیم شبکه ساده، تسریع سرعت یادگیری توسط نورون های لوکال و ویژگی های تقریب مناسب می شود.



همانطور که شکل 3 نشان می دهد رویکرد ما شامل دو ماژول مهم است: ماژول ایجاد چند پس زمینه و ماژول بررسی شی متحرک. ماژول ایجاد چند پس زمینه ارائه شده ، مدل پس زمینه احتمالی منعطفی است که به طورخودکار فاصله اقلیدسی را از هر پیکسل ورودی به کاندیدای پس زمینه مرجع محاسبه می کند. این امر به اطلاعات شبکه هایی بستگی دارد که در مراکز نورون لایه های پنهان قرار دارند.

یک مدل پس زمینه احتمالی می تواند محدوده پویایی از هر پیکسل را در پس زمینه مشخص کند و استفاده می شود تا بتوان لایه پنهان ساختار شبکه RBF را ایجاد کرد.

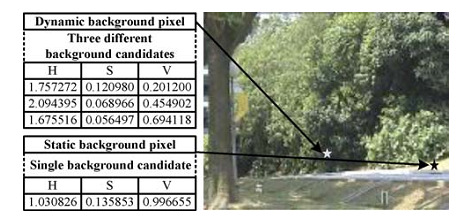
پس از بکار گیری ماژول MBG ، ماژول بررسی شی متحرک MOD استفاده می شود. ماژول MOD بررسی کامل و دقیقی از اشیا متحرک دارد که از دو رویه استفاده می کند: رویه هشدار بلوکی و رویه استخراج شی. رویه هشدار بلوکی آزمون غیر ضروری از منطقه پس زمینه پویا و ایستا را حذف میکند، که پس از آن فرایند استخراج شی این بلوک ها را پردازش کرده واحتمال بالای دسترسی به اشیا متحرک را خواهد داشت.

1. ایجاد چند پس زمینه

در این بخش، سه متغیر ادراکی HUE، اشباع و ارزش به عنوان لایه های ورودی وارد فضای رنگی HSV می شوند که بسیار مشابه با قابلیت دیداری انسانی است. فرض کنید که  نشان دهندهHUE، اشباع و ارزش باشند که به عنوان پیکسل  در هر فریم ورودی  دیده می شوند.

تعداد کافی نورون های پنهان می تواند دقت را بالا برد. با این وجود بسیاری از نورون ها ممکن است موجب افزایش ساختار شبکه و کاهش کیفیت عملکرد شوند. بنابراین مهم است تا مدل پس زمینه احتمالی منعطف مناسبی را ایجاد کرد که نشان دهنده نورون های پنهان باشد

به منظور ایجاد یک مدل پس زمینه احتمالی منعطف مناسب، هر ارزش شدت پیکسل ورودی  از t امین فریم  با نمونه های متناطر ارزشهای شدت پس زمینه  مقایسه می شوند. اگر شدت پیکسل ورودی  نزدیک به کاندیداهای شدت پس زمینه مرتبط باشد- برای مثال اگر پیکسل ورودی به کاندیدای پس زمینه تعلق داشته باشد – ما کاندیدای پس زمینه مرتبط با به روز می کنیم. از سوی دیگر  به عنوان یک کاندیدای جدید پس زمینه مشخص می شود



برای تعیین آنکه آیا پیکسل های ورودی  مرتبط با کاندیدای مرتبط با شدت پس زمینه هستند یا خیر، ما فاصله قلیدسی از بردارهای hexcone رنگی HSV استفاده میکنیم . این امر با استفاده از فاصله پیکسل  تا پیکسل محاسبه می شود



استفاده از این شاخص می تواند مساله را در دوره زمانی hue و و عدم ثبات hue برای ارزشهای اشباع کوچکتر حل کند.

یک تلورانس عملی  استفاده شده است تا تعیین کند که آیا پیکسل های ورودی  به کاندیدای پس زمینه تعلق دارند یا خیر در جاییکه  است. این قاعده تصمیم می تواند به شکل زیر بیان شود:



کاندیدای پس زمینه نزدیک به پیکسل ورودی  است که به شکل زیر است:



درجاییکه ، k امین کاندیدای اصلی و به روز شده در واقعیت  هستند و  یک پارامتر از قبل تعیین شده است. شکل 4 نشان دهنده شدت پس زمینه در محیط های پویا و ایستا است. این رویکرد ساختاری پس زمینه می تواند به عنوان یک فرایند یادگیری غیر نظارتی از مکان های مراکز در شبکه RBF مد نظر قرار گیرد

1. بررسی اشیا متحرک

1 . رویه هشدار بلوکی . ساختار شبکه RBF در اینجا شامل سه نورون وردی، یک نورون خروجی و یک لایه پنهان از نورون M است. ماژول MBG تعیین کننده تعداد M و نقاط مرکزی  از نورونهای لایه پنهان در شبکه RBF است که در شکل 2 نشان داده شده است. این امر همچنین نشان دهنده ساختار این شبکه است. پس از تعیین ساختار، عناصر HSV از پیکسل ورودی  به عنوان بردار ورودی استفاده می شوند. نورون های ورودی ، بردار ورودی نورون های لایه پنهان هستند. پس از محاسبه فاصله اقلیدسی میان بردار ورودی و نقاط مرکزی نورونن های پنهان ، خروجی نورون پنهان توسط تابع مبنا به شکل زیر بدست می آید:



در جاییکه  تابع مبناست ،  نقطه مرکزی I امین نورون است و p بردار ورودی است و m تعداد نورون های پنهان است و  فاصله اقلیدسی میان p و  است.

انواع مختلفی از توابع مبنا به طور مشابه استفاده شده اند همچن تابع گوسین، تابع خطی ، تابع کوبیک تبع thin

plate spline و غیره. برای رویکرد ما، از تابع مبنای رایج یعنی تابع گوسین استفاده می کنیم. تابع معرف به شکل زیر است:



در جاییکه به عنوان  تعریف می شود و  تلورانس تجربی فاصله اقلیدسی در 22 است. دلیل این امر کاهش  است که مرتبط با ایجاد یک کاندیدای پس زمینه در مدل احتمالی است. انحراف استاندار پایین تر ارزش  می تواند منحنی گوسین را شیب دارتر کند. این امر در ادامه از مجموع لایه خروجی از نسبت های  بدست می آید. بر اساس تجربیات ما می تواند به طور عملی به عنوان  تعریف شود.

بدلیل قابلیت عامل بندی و لوکال شدن تابع گوسین ، این تابع برای استفاده مناسب تر است. بعلاوه، تابع گوسین می تواند استفاده شود تا تناسب خوبی را ایجاد کند و وضعیت های بلوک ها را ارزیابی کند. ارزش خروجی بالاتر از تابع مبنا موجب نزدیک تر شدن بردار ورودی می شود که در نقاط مرکزی قرار گرفته است. برای مثال احتمال بالاتر پیکسل های ورودی پس زمینه. به منظور حذف آزمون ضروری منطقه پویا و ایستا، فریم ورودی به بلوکهای w\*W تقسیم می شود. محاسبه مجموع این توابع مبنا در هر بلوک به شکل زیر است:



در جاییکه p پیکسل مستقل از بلوک های متناظر  است، M تعداد نورون های پنهان و اندازه بلوک W است که به 4 نزدیک می شود.

وقتی مجموع بلوک های محاسبه شده  فراتر از آستانه S باشد، بلوک  صفر خواهد بود که نشان دهنده آن است که شامل پیکسل هایی نمی شود که به اشیا متحرک تعلق دارند. از سوی دیگر بلوک  با 1 نشان داده می شود که به معنی آن است بیشتر احتمال دارد که شامل پیکسل های اشیا متحرک باشد



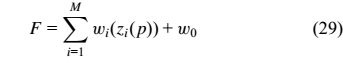
جدول 1 نشان دهنده مجموع توابع مبنا در بلوک ها در یک فریم ویدیویی ساده است. وقتی S معادل به 12 باشد، بلوک ها ممکن است شامل اشیا متحرکی باشند که باید بررسی شوند.

در نهایت، کاندیدای پس زمینه در لایه پنهان به شکل زیر خواهد بود:



درجاییکه  ، K امین کاندیدا در جایگاه  از مدلهای پس زمینه منعطف قبلی و فعلی است و پارامتر از قبل تعریف شده است. قاعده تصمیم که آیا  به تعلق دارد یا خیر بر اساس معادله 22 تعیین می شود

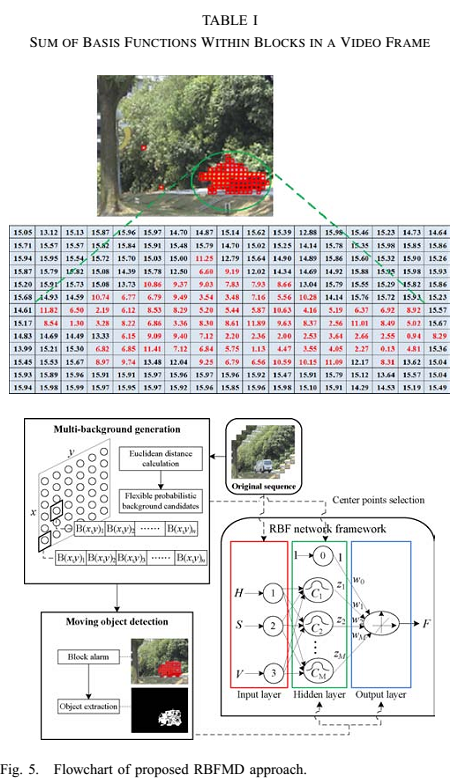
2 . رویه استخراج اشیا : پس از رویه هشدار بلوک ، آزمون های غیر ضروری حذف شده و رویه استخراج اشیا در بردارنده بلوک هایی است که شامل اشیا متحرک می شود. به عنوان آخرین گام این رویکرد، از لایه خروجی شبکه RBF، استفاده می شود تا حرکات باینری به عنوان نتیجه بررسی محاسبه شود. لایه خروجی تابع ترکیب خطی وزن دهی شده از ارزشهای پدید آمده از لایه پنهان را به شکل زیر محاسبه میکند:



در جاییکه  ، وزنی است که مرتبط با i امین نورون پنهان است و لایه خروجی  که ارزش خروجی از i امین نورون پنهان است و  آستانه ثابت است به طور اولیه ، مجموعه ای است که به طور عملی به 1 می رسد. کشف حرکات باینری همچون زیر است:



برچسب  با 1 ، به معنی آن است که پیکسل  بخشی از اشیا متحرک است. در غیر اینصورت  بخشی از پس زمینه است که با صفر نشان داده می شود. پس از اتمام عملیات در فریم های ورودی فعلی، وزن هایی برای عملیات به فریم بعدی داده می شود.



در شروع کار، همه وزن ها با 1 شروع می شود پس از آن وزن ها به شکل زیر تعدیل می شوند:



در جاییکه وزن میان لایه خروجی و i امین نورون پنهان در فریم  است،  نرخ یادگیری و M تعداد نورون های پنهان است.

پس از تعدیل وزن ها، وزن های ماین لایه های خروجی ونورون های پنهان که نزدیک به بردار ورودی است ، تقویت می شود و سایرین حذف می شوند. شکل 5 نشان دهنده فلوچارتی از رویکرد بررسی جرکت مبتنی بر RBF است.

5. نتایج عملی

تمایل این بخش به ارائه مقایسه ای میان روش RBFMD و چندین روش دیگر جدید است. نتایج عملی استخراج اشیا که توسط روش RBFMD انجام شده است از طریق مقایسات کیفی و کمی با سایر روشهای جدید این حوزه تحلیل شد تا توالی ویدیویی طبیعی برای محیط های پویا و ایستا نشان داده شود.

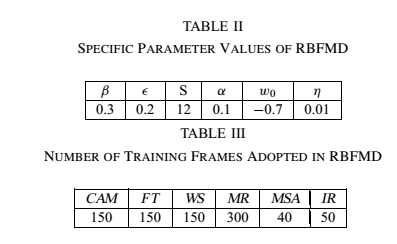
توالی CAM,WS,FT, و MR بکار گرفته شد و نتایج استخراج اشیا از پس زمینه های پویا آزمون شد. اولین توالی، CAM ، چرخ های متحرک و پیاده روی های جلوهای درختان در حال حرکت بودند. دومین توالی WS شخصی بود که روبروی فواره ای که آب از آن بالا می آمد حرکت می کرد. پس زمینه پویا در سومین توالی FT جریان آب خروشان بود و چهارمین توالی MR اتاق انتظاری بود که در آن پرده در حال حرکت بود.

دو توالی دیگر بکار برده شد تا نتایج بررسی روی پس زمینه های ایستا اندازه گیری شود. MSA توالی فردی بود که در راهرو قدم می زد و اشیا موقتا ثابت بودند. توالی IR مصاحبه شغلی بود که در آن یک فرد وارد می شد و به آرامی به گوشه ای حرکت می کرد.

همه این پارامتر ها در هر روش ، دارای ارزشهای بهینه هستند. بر اساس 36، تلورانس تعیین شده τ از روش RDBG می تواند در 5 قرار گیرد. برای حفظ پس زمینه احتمالی مدل با استفاده از GMM، تعداد عناصر گوسین K در 3 ثابت شدند که نرخ یادگیری α در آن معادل 0.001 بود. ادبیات روش SOBS نشان می دهد که نرخ یادگیری α از ثابت از قبل تعریف شده c1 و c2 بدست می آید که محدوده ای از 0.01 تا 1 را در بر می گیرد. بعلاوه آستانه  ، محدوده ای از 0.1 تا 0.1 و 0.01 تا 0.07 را در بر می گیرد. بر اساس 38، شفافیت حداقل و حداکثری همه پیکسل ها در روش CB مشتق شده از ارزشهای  و  بود که محدوده ای از 0.4 تا 0.7 را و محدوده ای از 1.1 تا 1.5 را در بر می گرفت.

در خصوص روش VIBE ، تعداد نمونه های پس زمینه هر N پیکسل می تواند 20 باشد، تعداد انطباق های زمانی  است که می تواند 2 باشد و تعیین کننده پیکسل پس زمینه باشد. آستانه از قبل تعریف شده فاصله اقلیدسی R در 20 قرار می گیرد. ارزشهای خاص همه پارامترها و تعداد فریم های آموزشی روش RBFMD در جدول 2 و3 نشان داده شده اند.

ما به طور متوالی الزامات حافظه را از مدل پس زمینه احتمالی اندازه گیری کردیم که توسط روش RBFMD مشخص شده بود و آن را با مدل های پس زمینه احتمالی سایر روشها مقایسه کردیم. در نهایت، ویژگی های محاسباتی برنامه های واقعی را تعیین کرده و سرعت پردازش روش RBFMD را اندازه گیری کردیم.



ارزیابی کیفی

به منظور ارزیابی هدفمند دقت باینری اشیا که توسط RBFMD و سایر روشهای جدید بیان شده است، ارزیابی های کیفی از طریق به یادآوری، دقت F1 و شاخص های شباهت تعیین می شوند که برای آزمون توالی های ویدیویی مناسب است

شاخص به یادآوری، نشان دهنده درصد واقعیت های درست بررسی شده توسط مقایسه ای است که در ان حساب کل آیتم ها در واقعیت به شکل زیر است:



در جاییکه tp میزان کل پیکسل های مثبت واقعی و fn میزان کل پیکسل های منفی اشتباه است و نشان دهنده کل آیتم های واقعیت زمینه است.

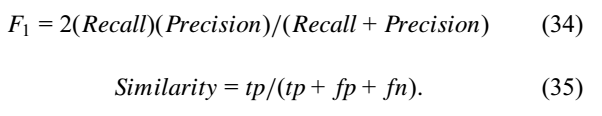
دقت نشان دهنده درصدی است که واقعیت های مثبت بررسی شده توسط مقایسه میزان کل آیتم ها در اهداف باینری توسط این روش بدست می آیند که به شکل زیر است:



در جاییکه fp میزان کل پیکسل های مثبت اشتباه و میزان کل آیتم های محاسبه شده در اشیا باینری است.

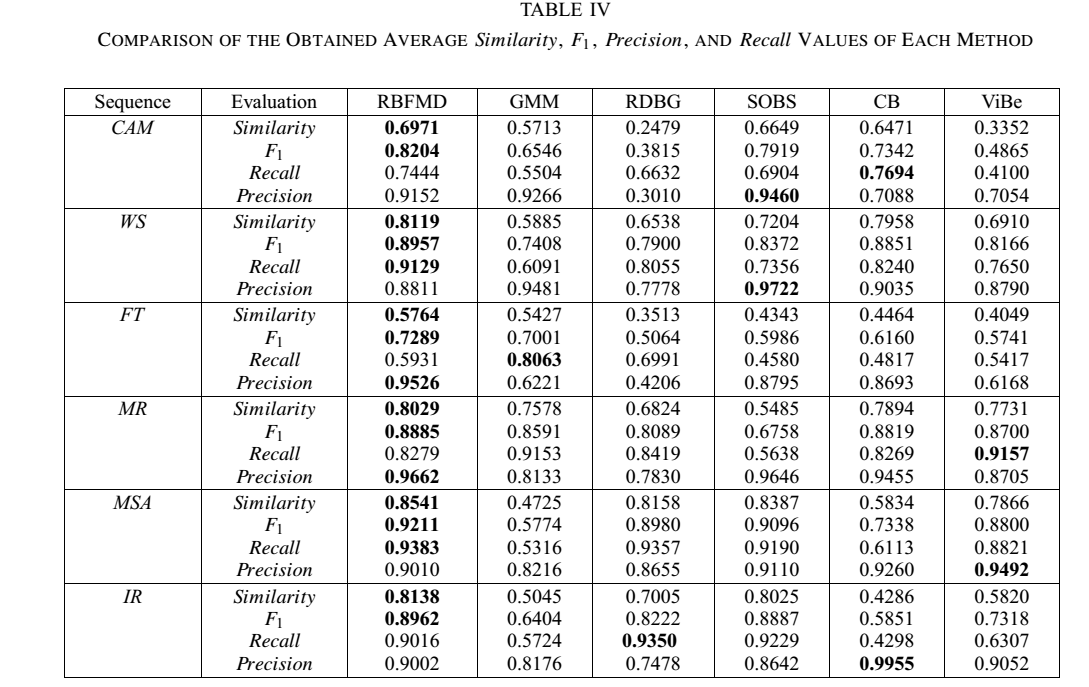
با این وجود، به یاد اوردن، به طور انتخابی فقط ارتباطات آیتم های اشتباه درونی از اشیا متحرک را اندازه گیری می کند و دقت، به طور انتخابی فقط ارتباطات اشتیابه از آیتم های بررسی شده را اندازه گیری می کند. بر این اساس، با استفاده از شاخص های ذکر شده بالا نمی توان مقایسه رضایت بخشی را میان روشهای مختلف انجام داد

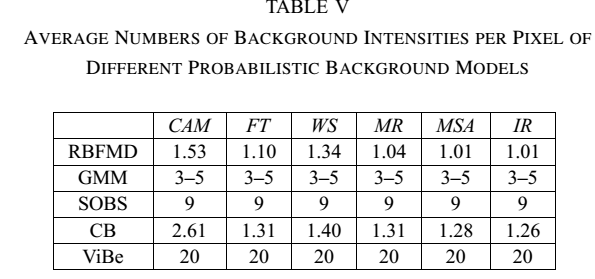
به منظور تسهیل یک اندازه گیری کارامد، دقت بر اساس دو شاخص دیگر F1 و شباهت ارزیابی می شوند. استفاده از این شاخص ها توسط وزن دهی میانگین هارمونیک یادآوری و دقت بدست می آید که :



همه ارزشهای بدست آمده از طریق شاخص های مد نظر، محدوده 0 تا 1 را در بردارد با ارزشهای بالاتری که به معنی دقت بیشتر است.

میانگین ارزشهای دقت برای آزمون همه توالی ها از طریق چهار شاخص فوق بدست می آید. اینها از روشهای GMM [29], RDBG [36], SOBS [37], CB [38], ViBe [39], and RBFMD بدست آمده اند و در جدول 4 نشان داده شده اند. ما میتوانیم مشاهده کنیم که روش RBFMD بهترین شباهت و ارزش F1 را در مقایسه با سایر روشهای جدیدی چون CAM, WS, FT, MR, MSA, and و توالی های IR دارد. به طور خاص این روش تنها روشی است که نرخ دقت همه شاخص هایش بیش از 80 درصد برای توالی های WS و MR است که در بردارنده پیش زمینه پویا هستند.

در مقایسه نرخ های دقت روش GMM که از طریق شباهت و F1 بدست آمده اند بیش از 38 تا 34 درصد ، پایین تر از آنهایی هستند که در روش RBFMD بدست آمده اند. نرخ های دقت بدست آمده از شاخص های شباهت و F1در روش RDBG بیش از 45 تا 44 درصد پایین تر از انهایی است که در روش RBFMD بدست آمده است . نرخ های دقت بدست آمده از طریق شباهت و F1 برای روش SOBS بیش از 25 و 21 درصد پایین تر از نرخ هایی است که در روش RBFMD بدست آمده است. نرخ دقت بدست آمده توسط شباهت و F1 در روش CB ، بیش از 39 و 31 درصد پایین تر از نرخ های بدست آمده در روش RBFMD است. نرخ دقت بدست آمده توسط شباهت و F1 در روش VIBE ، بیش از 36 و 33 درصد پایین تر از نرخ های بدست آمده در روش RBFMD است. مهم است تا یادآوری شود که هر روش پیاده شده توسط پارامتر های بهینه مطالعات قبلی و همه ارزشهای دقت هر مدل با توجه به هر فریم از آزمون توالی ها بدست آمده است.



2. ارزیابی کیفی

در اینجا ، ارزیابی کیفی از نتایج استخراج شده اشیا برای توالی آزمون های مختلف برای هر روش توسط بررسی دیداری انجام شده است. ارزیابی کیفی بدست آمده از هر توالی ویدیویی در کنار ارزشهای دقت شباهت و F1 از اشیا باینری در هر روش در شکل 6-9 نشان داده شده است. با مقایسه اشیا باینری ایجاد شده در واقعیت های زمینه ، به این نتیجه رسیده ایم که روش RBFMD بیشترین اکتشاف حرکتی را دارد نه تنها در صحنه های پویا بلکه همچنین در صحنه های استا و نتایج اکتشاف RBFMD بسیار دقیق تر از سایر روشهای GMM, RDBG, SOBS, CB, and ViBe است

مشخص است که نتایح ارزیابی های کمی و کیفی نشان می دهد که روش RBFMD توانسته است به طور موفقیت آمیزی حرکات اشیا را هم در پیش زمینه های پویا و هم در محیط های ایستا درک کند که عملکرد بالاتری از سایر روشهای موجود دارد.

1. تحلیل چند پس زمینه

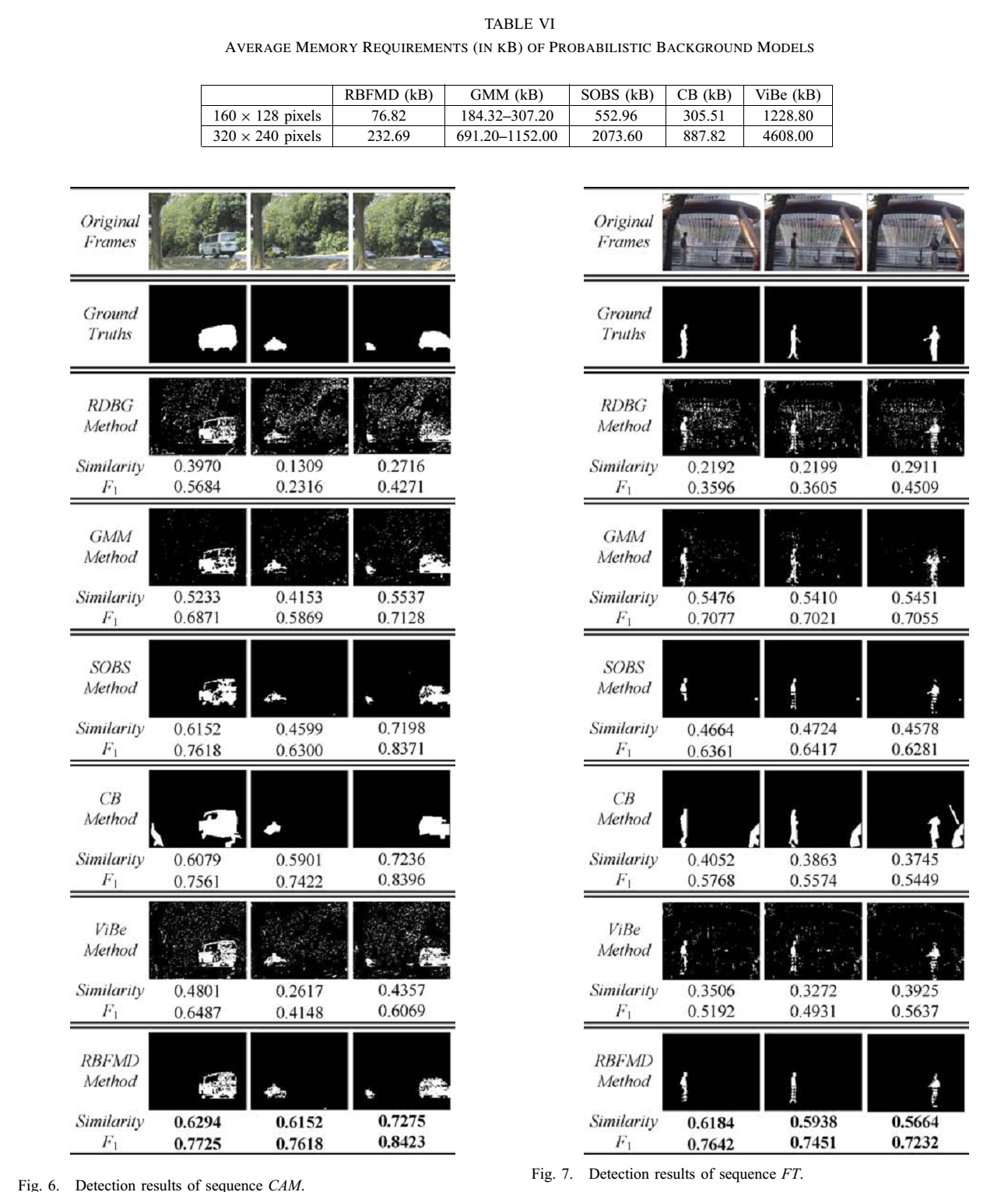
در سیستم های نظارتی ویدیویی پیشرفته، یک مدل پس زمینه مناسب برای اکتشاف دقیق موردنیاز است. اکثر رویکردها از یک تصویر ساده برای مشخص کردن پس زمینه استفاده می کنند، در عین حال تلاش می کنند تا اکتشاف دقیق تری نیز داشته باشند. با این حال، معمولا دشوار است تا تخمین های درستی از پیش زمینه های پویا زد و از مدل های پس زمینه ساده استفاده کرد. در کل، مدل پس زمینه احتمالی برای درک پس زمینه پویا بسیار مناسب تر است. با این حال، استفاده از مدل های پس زمینه احتمالی ممکمن است الزامات حافظه ای را افزایش دهد. مدل پس زمینه احتمالی منعطف تعداد زیادی از شدت های پس زمینه را در هر موقعیت پیکسلی با توجه به محدوده پویای هر پیکسل ذخیره می کند.

جدول 5 نشاندهنده میانگین تعداد شدت هر پیکسل از مدل های پس زمینه احتمالی متفاوت است که توسط RBFMD, GMM, SOBS, CB و ViBe ایجاد شده است و توالی آزمون ها روی آن انجام شده

ما میتوانیم ببینیم که میانگین تعداد کاندیداهای پس زمینه ذخیره شده برای هر پیکسل با استفاده از روش RBFMD برای توالی شامل پس زمینه پویا می شود. (CAM, FT, WS, و MR) که 1.53 یا کمتر است و تعداد توالی برای پس زمینه ایستا (MSAand IR)نزدیک به 1 است. در مقایسه GMM, SOBS, و ViBe تعداد شدت های پس زمینه در هر موقیت پیکسلی است. در مطالعات قبلی ، اعداد ذخیره شده GMM, SOBS و ViBe به ترتیب 3، 5، 9 و 20 پیشنهاد شده اند

میانگین الزامات حافظه مورد نیاز برای تصاویر اندازه های مختلف برای مدل های پس زمینه احتمالی روشهای مختلف در جدول 6 نشانداده شده است. آزمون توالی های (CAM, FT, WS, و MR) با تصویری به اندازه 160 در 128پیکسل و توالی MSA و IR با تصویری با اندازه 320 در 240 پیکسل بکار گرفته شد تا الزامات میانگین حافظه ارزیابی شود.

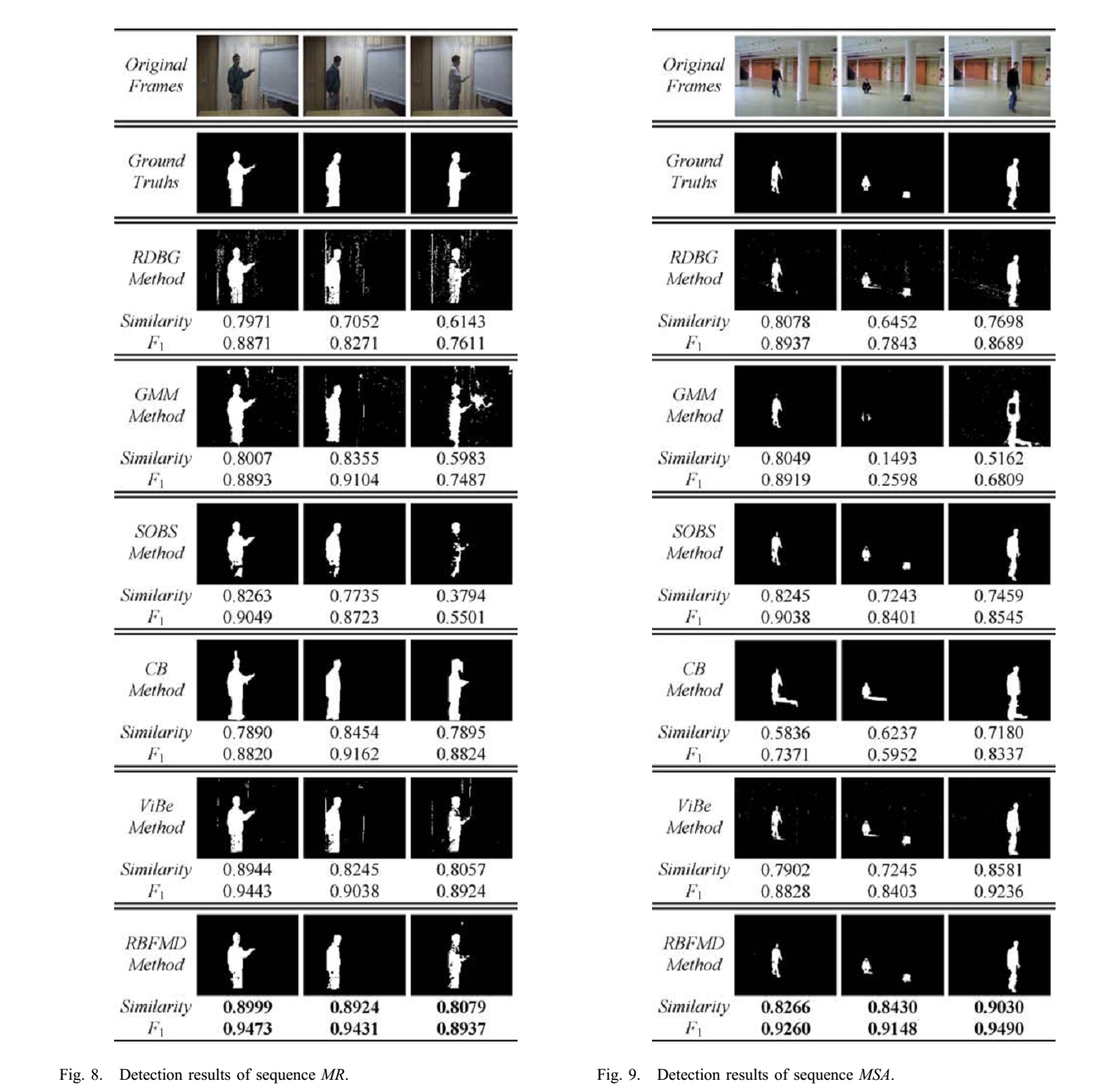
برای توالی تصویری بااندازه 160 در 128 ، میانگین حافظه مورد نیاز در روش RBFMD حدودا 76.82 Kb است. در حالیکه برای تصویری با ندازه 320 در 240 پیکسل ، 232.69 kB مورد نیاز است. در مقایسه با الزمات حافظه ای برای مدل های پس زمینه احتمالی ازسایر روشها، استفاده از روش ارائه شده می تواند موجب کاهش 58 تا 95 درصدی در حافظه مورد نیاز شود.

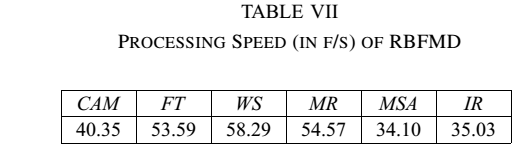


**نتایج عملکرد**

برای تایید امکان سنجی محاسباتی برای کاربردهای واقعی، ما سرعت پردازش روشRBFMD را با چند توالی آزمون در جدول 7 گزارش دادیم. روش RBFMD با استفاده از زبان برنامه نویسی C روی پردازشگر Core2Quad 2.33 GHz با 2 GB of RAM و با سیستم عامل ویندوز 7 انجام شد.

نتایج عملکرد نشان می دهد که برای همه توالی ها با پس زمینه های پویا و ایستا، روش RBFMD می تواند سرعتی بالاتر از 34 f/s را داشته باشد که برای انجام برنامه های واقعی کافی است.





**5. نتیجه گیری**

یک رویکرد کشف حرکات جدید برای بخش بندی اشیا، در محیط های پویا و ایستا ارائه شد. این روش ارائه شده بر اساس شبکه طبیعی RBF است و ترکیبی از ماژول های ایجاد چند پس زمینه ای منحصر به فرد را در کنار ماژول بررسی اشیا متحرک دو رویه ای جدید قرار می دهد تا بتوان به بررسی های دقیق و کاملی در محیط های پویا و ایستا دست پیدا کرد. ماژول MBG به طور کارامدی مدل احتمالی منعطفی را شکل می دهد که می تواند بیانگر محدوده پویایی از هر پیکسل در پس زمینه باشد و میتواند برای شکل دادن به لایه های پنهان ساختار شبکه RBF بکار رود. این امر ساختار شبکه را حداقل کرده و سرعت پردازش راافزایش می دهد. پس از ایجاد مدل پس زمینه احتمالی با کیفیت، ماژول MOD ازرویه هشدار بلوکی استفاده میکنند تا آزمون های غیر ضروری را از کل منطقه پس زمینه حذف کند و پس از ان رویه استخراج شی می تواند به طور کارامدی پیکسل های اشیا متحرک را بررسی کند. برای مرحله نهایی، وزن خروجی ها برای عملیات روی فریم های ورودی بعدی تعدیل شد. نتایج عملی توسط مقایسات کمی و کیفی با سایر روشهای جدید بر اساس محدوده وسیعی از توالی های ویدویی طبیعی انجام گرفت. نتایج ارزیابی های کمی و کیفی نشان داد که روش ارائه شده قادر است تا بررسی های دقیق وکافی را هم در محیط های ایستا و هم در محیط های پویا داشته باشد و بهتر از روشهای دیگر عمل کند. بعلاوه، ما به طور عملی نشان دادیم که روش ارائه شده الزامات حافظه ای حداقلی در ایجاد مدل پس زمینه احتمالی دارد و می تواند در مقایسه با سایر تکنیک های مدلسازی پس زمینه احتمالی استفاه شود. بعلاوه روش ما همچنین برای برنامه های واقعی نیز کارامد است. تحلیل ها نشان می دهد که روش RBFMD ارائه شده می تواند اشیا متحرک را هم در محیط های ایستا و هم در محیط های پویا با محاسبات ارزان قیمت و الزامات حافظه ای پایین انجام دهد.