



دانشکده مهندسی برق و رباتیک

پایان نامه کارشناسی ارشد

دنبال کردن شیء در رشته ویدئویی به کمک توصیفگرهای آماری

امیر نبئی

اساتید راهنما:

دکتر علیرضا احمدی فرد

دکتر علی سلیمانی

بہمن ۸۷

بسم اللہ الرحمن الرحیم



الف



دانشگاه صنعتی شاهرود

دانشکده مهندسی برق و رباتیک

گروه الکترونیک

دنبال کردن شیء در رشته ویدئویی

به کمک توصیفگرهای آماری

دانشجو: امیر نبئی

اساتید راهنما:

دکتر علیرضا احمدی فرد

دکتر علی سلیمانی

استاد مشاور:

دکتر حسین مروی

پایان نامه کارشناسی ارشد جهت اخذ درجه کارشناسی ارشد

بهمن ۸۷

ب

شماره: ۵۳۲۲/۰۶
تاریخ: ۱۵، ۱۱، ۸۷
ویرایش: —



مدیریت تحصیلات تکمیلی
فرم شماره (۶)

بسمه تعالی

فرم صورتجلسه دفاع پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای امیرمحمدی رشته مهندسی عمران گرایش سازه تحت عنوان بررسی سازه های فولاد-بتن در مناطق زلزله خیز که در تاریخ ۸۷/۱۱/۱۴ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح زیر است:

قبول (با درجه: عالی - امتیاز ۱۸/۷۰) ☐ دفاع مجدد ☐ مردود ☐

۱- عالی (۲۰ - ۱۸) ۲- بسیار خوب (۹۹/۱۷ - ۱۶)

۳- خوب (۹۹/۱۵ - ۱۴) ۴- قابل قبول (۹۹/۱۳ - ۱۲)

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنما	عبدالله احمدی	استاد ارشد	
۲- استاد مشاور	محمد حسن	استاد ارشد	
۳- نماینده شورای تحصیلات تکمیلی	محمدحسین زعفرانی	استاد ارشد	
۴- استاد ممحن	آقای دکتر محمدی	استاد ارشد	
۵- استاد ممحن	اسدالله محمودی	استاد ارشد	

تأیید رئیس دانشکده:

تعهد نامه

اینجانب **امیر نبی** دانشجوی دوره کارشناسی ارشد / دکتری رشته **برق - الکترونیک**
 دانشکده **مهندسی برق** دانشگاه صنعتی شاهرود نویسنده پایان نامه / رساله **دینامیک کربن شیء در شبکه های برق**
 بر کمک **دکتر سید علی آقامری** تحت راهنمایی **دکتر احمد زین الدین سیفی** متعهد می شوم .

- تحقیقات در این پایان نامه / رساله توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه / رساله تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه / رساله تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه / رساله رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه / رساله ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه / رساله ، 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: ۸۸/۱/۱۸
 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 است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان نامه / رساله بدون ذکر مرجع مجاز نمی باشد.

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

مسئله دنبال کردن شیء در یک رشته ویدئویی یکی از مسائل مهم در شاخه ماشین بینایی محسوب می شود. هدف الگوریتمهای دنبال کردن شیء یافتن موقعیت شیء در فریمهای یک رشته ویدئویی می باشد و در بعضی از موارد هم این هدف شامل بدست آوردن اطلاعات دیگری از شیء مانند محدوده ناحیه ای که توسط شیء اشغال شده است نیز می شود.

در یک سیستم دنبال کننده شیء یکی از مسائل مهم، طریقه بیان شیء می باشد. در این تحقیق روش جدیدی برای بیان شیء در دنبال کننده مبتنی بر انتقال میانگین پیشنهاد می گردد. در روش پیشنهادی علاوه بر اطلاعات رنگ که استفاده از آن در دنبال کننده مبتنی بر انتقال میانگین متداول است، اطلاعات جهت و دامنه گرادیان را نیز به کار گرفته ایم. در بخش نتایج نشان می دهیم که چنانچه رنگهای به کار رفته در پس زمینه و شیء متحرک مشابه باشند دنبال کننده مبتنی بر اطلاعات رنگ ناموفق خواهد بود در حالی که روش پیشنهادی با موفقیت شیء را دنبال می کند.

دنبال کننده دیگری که مورد توجه بسیاری از محققین قرار گرفته است، دنبال کننده مبتنی بر فیلتر ذره ای می باشد. در بخش بعدی این تحقیق، جهت بهبود سرعت و دقت عملکرد دنبال کننده مبتنی بر فیلتر ذره ای، پیشنهاد می کنیم که بر روی بعضی از ذره ها، الگوریتم انتقال میانگین اعمال شود. نتایج آزمایشات انجام شده نشان میدهد با استفاده از این روش می توان به دنبال کننده سریعتر و دقیقتری نسبت به دنبال کننده مبتنی بر فیلتر ذره ای متداول دست یافت. مقایسه این الگوریتم پیشنهادی با دنبال کننده انتقال میانگین هم نشان دهنده این مسئله است که دنبال کننده پیشنهادی در مواردی که دنبال کننده انتقال میانگین شیء را در اثر سرعت زیاد جابجایی آن گم می کند، قادر به دنبال کردن شیء می باشد ضمن اینکه ویژگی مطلوب دنبال کننده انتقال میانگین یعنی سرعت بالای اجرای آن را تا حدود زیادی حفظ می کند.

کلمات کلیدی: دنبال کردن شیء- انتقال میانگین - فیلتر ذره ای- هیستوگرام جهت گرادیان

فهرست

فصل اول: معرفی ۱

۱-۱ مقدمه ۲

۲-۱	تعریف مسئله	۲
۳-۱	مشکلات سیستم دنبال کننده شیء	۳
۴-۱	کاربردها	۴
۵-۱	ساختار پایان نامه	۴
۶	فصل دوم: مروری بر کارهای انجام شده در زمینه دنبال کردن شیء	۶
۱-۲	مقدمه	۷
۲-۲	بیان شیء	۷
۳-۲	انتخاب ویژگی برای دنبال کردن شیء	۱۲
۴-۲	آشکار سازی شیء	۱۵
۱-۴-۲	آشکار سازی نقاط مهم	۱۶
۲-۴-۲	تفریق پس زمینه	۱۶
۳-۴-۲	بخش بندی تصویر	۱۷
۴-۴-۲	یادگیری با سرپرست	۱۸
۵-۲	دنبال کردن شیء	۱۹
۱-۵-۲	دنبال کردن نقطه	۲۰
۱-۵-۲-۱	روشهای معین برای تعیین تناظر	۲۱
۲-۵-۲-۱	روشهای آماری برای تعیین تناظر	۲۳
۲-۵-۲	دنبال کردن کرنل	۲۵
۱-۵-۲-۱	دنبال کردن با استفاده از الگوها و مدل‌های مبتنی بر تابع چگالی ویژگیهای ظاهر شیء	۲۶
۲-۵-۲-۲	دنبال کردن با استفاده از مدل‌های ظاهر چند نمایی	۲۷
۳-۵-۲	دنبال کردن سایه نما	۲۹
۱-۳-۵-۲	تطبیق شکل	۳۰
۲-۳-۵-۲	دنبال کردن کانتور	۳۲
۶-۲	مسائل مرتبط	۳۶
۱-۶-۲	حل مسئله مخفی شدن شیء	۳۶

۳۷	۲-۶-۲ دنبال کردن با استفاده از چند دوربین
۳۹	فصل سوم: تئوری
۴۰	۳-۱ مقدمه
۴۱	۳-۲ شیوه انتقال میانگین
۴۷	۳-۳ دنبال کردن شیء با استفاده از روش انتقال میانگین
۵۴	۳-۴ فیلتر کالمن
۶۰	۳-۵ فیلتر ذره ای
۶۵	۳-۶ دنبال کننده TSHT
۶۶	فصل چهارم: روشهای پیشنهادی
۶۷	۴-۱ استفاده از اطلاعات جهت و دامنه گرادیان در دنبال کننده انتقال میانگین
۷۳	۴-۲ استفاده از روش انتقال میانگین در فیلتر ذره ای
۷۶	فصل پنجم: نتایج عملی
۷۷	۵-۱ استفاده از اطلاعات جهت و دامنه گرادیان در دنبال کننده انتقال میانگین
۸۱	۵-۲ استفاده از روش انتقال میانگین در فیلتر ذره ای
۹۳	فصل ششم: خلاصه و نتیجه گیری
۹۸	فهرست مراجع

فهرست اشکال

شیء	مختلف	بیانهای	شکل ۱-۲
	۹.....		
ویدئوی	رشته	یک	شکل ۲-۲
		از	تصویری
		۱۴.....	
نوری		های	کل ۳-۲
		جریان	
	۱۴.....		
شده	آشکارسازی	مهم	شکل ۴-۲
		نقاط	
	۱۶.....		
۱۷.....			شکل ۵-۲ آشکارسازی با استفاده از تفریق پس زمینه
۱۸.....			شکل ۶-۲ بخش بندی تصویر با استفاده از روش انتقال میانگین
۲۰.....			شکل ۷-۲ دسته بندی روشهای دنبال کردن شیء
نقاط	بین	تناظر	شکل ۸-۲
		برقراری	
	۲۱.....		
مختلف		حرکتی	شکل ۹-۲
		قیود	
	۲۲.....		
شیء	نمای	سایه	شکل ۱۰-۲
	۳۰.....		
.....			شکل ۱-۳ روند نمای فیلتر کالمن
			۵۹
.....			شکل ۲-۳ عملکرد الگوریتم انتشار چگالی شرطی
			۶۴
۶۸			شکل ۱-۴ شیء در پس زمینه ای با رنگهای مشابه
.....			شکل ۲-۴ هیستوگرام جهت گرادیان
			۶۸
۷۱.....			شکل ۳-۴ تصویر شیء و بخشی از پس زمینه با رنگهای مشابه
هیستوگرامها	مقایسه	۴-۴	شکل
	۷۲.....		

- شکل ۵-۱ اعمال روش انتقال میانگین بر روی رشته ۱ با استفاده از هیستوگرام رنگ..... ۷۸
- شکل ۵-۲ اعمال روش انتقال میانگین بر روی رشته ۱ با استفاده از هیستوگرام پیشنهادی ۷۸
- شکل ۵-۳: اعمال روش انتقال میانگین بر روی رشته ۲ با استفاده از هیستوگرام رنگ ۷۹
- شکل ۵-۴: اعمال روش انتقال میانگین بر روی رشته ۲ با استفاده از هیستوگرام پیشنهادی ۷۹
- شکل ۵-۵: خطای دنبال کننده و مشابهت بدست آمده ۸۰
- شکل ۵-۶: فریم ح های ۵،۴،۳ و ۶
.....
- شکل ۵-۷ مقایسه خطا و سرعت جسم
۸۵
- شکل ۵-۸ جابجایی یک شیء از یک فریم به فریم دیگر ۸۸
- شکل ۵-۹ اشیائی با ابعاد مختلف و جابجایی های یکسان ۸۹
- شکل ۵-۱۰ تاثیر میزان جابجایی شیء در سرعت اجرای الگوریتم پیشنهادی ۹۰
- شکل ۵-۱۱ تعداد تکرارهای انتقال میانگین در هر فریم در مقایسه با سرعت جسم ۹۲

فهرست جداول

ط

جدول ۵-۱: مقایسه روشهای TSHT ، دنبال کننده مبتنی بر فیلتر ذره ای و دنبال کننده مبتنی بر انتقال میانگین با روش پیشنهادی ۸۶

فصل اول

معرفی

۱-۱ مقدمه

امروزه مسئله دنبال کردن شیء در یک رشته ویدئویی یکی از مسائل مهم و پرکاربرد در شاخه ماشین بینایی می باشد. ظهور کامپیوترهای قدرتمند، در دسترس بودن دوربینهای ارزان قیمت و نیاز فزاینده به تحلیل ویدئویی خودکار، علاقه زیادی را جهت توسعه الگوریتمهای دنبال کردن شیء ایجاد کرده است.

۱-۲ تعریف مسئله

دنبال کردن شیء را می توان به صورت پیدا کردن موقعیت شیء مورد نظر در تک تک فریمهای یک رشته ویدئویی در نظر گرفت. علاوه براین در بعضی از کاربردها، سیستم دنبال کننده شیء ممکن است اطلاعات دیگری مانند شکل شیء را نیز معلوم کند. برای مثال در یک مسابقه فوتبال شیء متحرک مورد نظر می تواند یک توپ باشد که سیستم دنبال کننده باید موقعیت آن را در هر فریم در اختیار ما قرار دهد و یا شیء مورد نظر می تواند یک انسان در حال حمل کیف باشد، در این صورت ممکن است از سیستم دنبال کننده این انتظار را داشته باشیم که با تعیین ناحیه کامل شیء در هر فریم تغییرات آن را نیز ثبت کند مثلاً در صورتی که فرد کیف را رها کند یا به شخص دیگری تحویل دهد، دنبال کننده متوجه این تغییر و جابجایی شیء بشود و اطلاعات مربوطه را در اختیار بخشهای بعدی سیستم قرار دهد.

۱-۳ مشکلات سیستم دنبال کننده شیء

دنبال کردن شیء در حالت کلی مسئله ای دشوار است . مشکلات موجود در این زمینه می تواند ناشی از

موارد زیر باشد:

- از دست دادن اطلاعات در اثر تصویر کردن دنیای سه بعدی در یک تصویر دو بعدی

- نویز در تصویر

- حرکت پیچیده شیء

- ماهیت غیر صلب یا مفصلی^۱ اشیاء

- مخفی شدن شیء^۲ به صورت کلی یا جزئی

- پیچیدگی شکل شیء

- تغییرات در شدت نور صحنه

- نیاز به پردازش سریع (بلادرنگ^۳).

برای ساده کردن مسئله دنبال کردن شیء می توان قیودی را برای ظاهر و/یا حرکت شیء در نظر گرفت

برای مثال بسیاری از الگوریتمهای دنبال کردن شیء فرض می کنند که حرکت شیء تغییرات ناگهانی

¹ Articulated

² Occlusion

³ Real-time

نخواهد داشت یا در بعضی از این الگوریتمها فرض می شود که حرکت شیء دارای سرعت یا شتاب ثابت است. اطلاعات اولیه در باره اندازه یا ظاهر شیء نیز می تواند برای ساده سازی مسئله به کار گرفته شود.

۴-۱ کاربردها

امروزه دنبال کردن شیء کاربرد گسترده ای در زمینه های مختلف پیدا کرده است از جمله این کاربردها می توان به موارد زیر اشاره کرد:

- بازشناسی بر اساس حرکت مانند شناسایی افراد از طریق طرز راه رفتن

- نظارت خودکار

- ارتباط انسان-کامپیوتر مانند شناسایی اشارات^۴

- ناوبری وسایل نقلیه

۵-۱ ساختار پایان نامه

در فصل دوم مروری بر کارهای انجام شده در زمینه دنبال کردن شیء صورت گرفته است. با توجه به اینکه در دنبال کردن شیء نیازمند روشی برای بیان شیء هستیم و همچنین باید ویژگی مناسب را برای این کار انتخاب کنیم، در فصل دوم ابتدا به بررسی روشهای مختلف برای بیان شیء خواهیم پرداخت و سپس ویژگی های متداول را مطرح خواهیم کرد و پس از آن آشکارسازی شیء را که در بسیاری از کاربردهای دنبال کردن

⁴ Gesture

شیء مورد نیاز است مطرح خواهیم کرد و به روشهای موجود در این زمینه می پردازیم . پس از آن به سراغ روشهای دنبال کردن شیء می رویم و آنها را در سه دسته دنبال کردن نقطه ، دنبال کردن کرنل و دنبال کردن سایه نما طبقه بندی می کنیم. در فصل سوم به چند روش در زمینه دنبال کردن شیء به طور دقیق تر می پردازیم این روشها شامل انتقال میانگین [۱و۲] ، فیلتر کالمن [۳]، فیلتر ذره‌ای [۴و۵] و TSHT^۵ [۶] می باشند. در فصل چهارم روشهای پیشنهادی خود را مطرح می کنیم. اولین روش در این بخش روشی برای پیاده سازی الگوریتم دنبال کننده انتقال میانگین است که در آن از اطلاعات جهت و دامنه گرادیان استفاده می‌شود. و دومین روش ، یک روش دنبال کردن شیء مبتنی بر فیلتر ذره‌ای است که در آن از الگوریتم انتقال میانگین استفاده شده است. فصل پنجم شامل بررسی نتایج پیاده سازی روشهای پیشنهادی و مقایسه آنها با برخی از روشهای موجود می باشد و فصل ششم به جمع بندی و نتیجه گیری اختصاص دارد.

⁵ Two Stage Hybrid Tracker

فصل دوم

مروری بر کارهای انجام شده
در زمینه دنبال کردن شیء

۲-۱ مقدمه

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 می‌بریم . اولین مسئله تعریف یک بیان^۶ مناسب از تصویر شیء متحرک است و مسئله بعدی انتخاب ویژگی‌های تصویر می باشد که به عنوان ورودی دنبال کننده استفاده می‌شوند. بنابر این در این بخش ابتدا به مرور روشهای مختلف بیان شیء و پس از آن به ویژگی های متداولی که برای دنبال کردن شیء قابل به کار گیری هستند می پردازیم . تقریباً تمام الگوریتمهای دنبال کردن نیاز به آشکارسازی شیء در فریم اول یا در همه فریمها دارند[۷]. در قسمت بعدی این فصل نگاهی به روشهای آشکار سازی شیء خواهیم داشت و بعد از آن به روشهای موجود برای دنبال کردن شیء می پردازیم.

۲-۲ بیان شیء

⁶ Representation

در یک الگوریتم دنبال کردن شیء، در ابتدا باید شیء را به الگوریتم معرفی کنیم یا به عبارت دیگر شیء را بیان کنیم. اشیاء را می‌توان با شکل و ظاهر آنها بیان کرد. در این بخش ما در ابتدا بیانهای شکل شیء را که استفاده از آنها در الگوریتمهای دنبال کردن شیء متداول هستند تشریح خواهیم کرد و سپس به بیانهای شکل و ظاهر در کنار هم، می‌پردازیم.

نقطه‌ها: شکل را می‌توان با یک نقطه (شکل ۲-۱ الف) [۸] یا با مجموعه‌ای از نقاط (شکل ۲-۱ ب) [۹] بیان نمود. در حالت کلی بیان با نقطه برای دنبال کردن اشیائی که ناحیه کوچکی از تصویر را اشغال می‌کند مناسب می‌باشد.

اشکال هندسی پایه: شکل شیء با یکی از اشکال هندسی پایه مانند مستطیل (شکل ۲-۱ ج) یا بیضی (۲-۱ د) [۲] قابل بیان کردن می‌باشد. حرکت شیء برای چنین بیانیهایی معمولاً با تبدیلهای هندسی پروجکتیو^۷ (هوموگرافی)، انتقال یا افاین^۸، مدل می‌شود. گرچه اشکال هندسی اولیه برای بیان اشیاء صلب مناسب‌تر هستند، لیکن برای دنبال کردن اشیاء غیرصلب نیز به کار می‌روند.

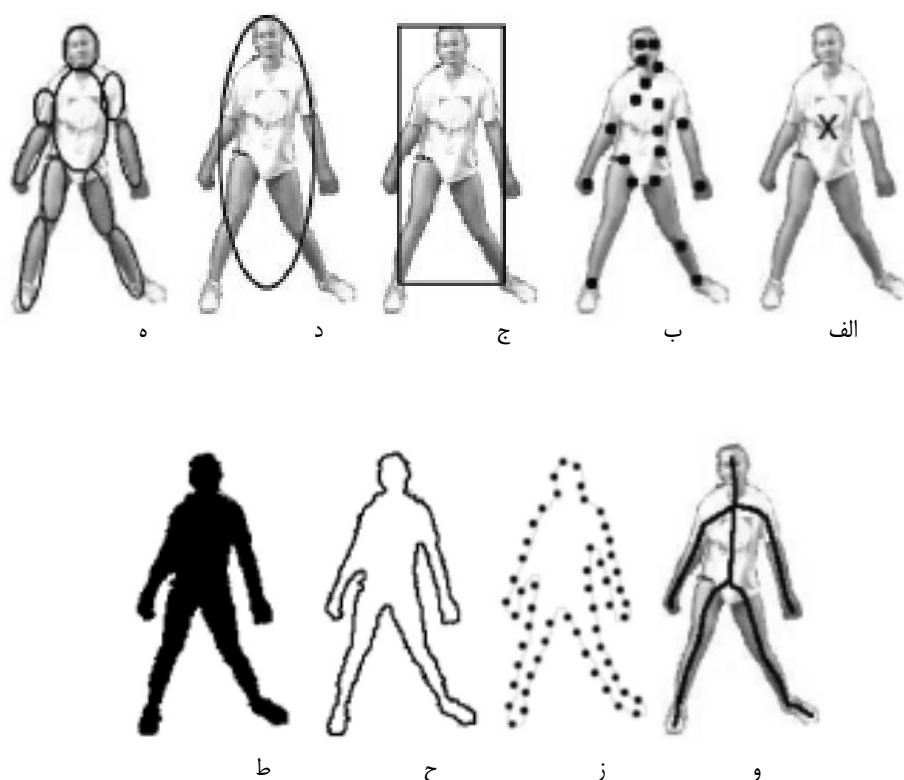
⁷ Projective

⁸ Affine

کانتور^۹ و سایه‌نمای^{۱۰} شیء: بیان کانتور، مرز شیء را تعریف می‌کند. (شکل ۱-۲ ز و ۱-۲ ح). ناحیه

درون مرز، سایه‌نمای شیء نامیده می‌شود (شکل ۱-۲ ط). بیان کانتور و سایه‌نما برای دنبال کردن اشیاء

پیچیده ی غیر صلب مناسب هستند. [۱۰].



شکل ۱-۲ بیانهای شیء: الف) یک نقطه ب) چند نقطه، ج) ناحیه مستطیلی، د) ناحیه بیضوی ه) نواحی چندگانه که به یکدیگر متصل هستند و) اسکلت شیء ز و ح) کانتور شیء ط) سایه نما

⁹ Contour

¹⁰ Silhouette

مدلهای شکل بندبند^{۱۱}: اشیاء بندبند از قسمتهایی تشکیل شده‌اند که با اتصالاتی به هم متصل می‌شوند. برای مثال بدن انسان یک شکل بندبند است با تنه، پاها، دستها و سر که با اتصالاتی در کنار هم قرار گرفته‌اند برای بیان یک شیء بندبند می‌توان اجزاء تشکیل دهنده را با استوانه‌ها یا بیضی‌هایی مدل کرد (شکل ۲-۵۱).

مدلهای اسکلتی: اسکلت شیء را می‌توان با اعمال تبدیل medial axis بر روی سایه نمای شیء، استخراج کرد [۱۱]. این مدل معمولاً در بیان شکل برای بازشناسی اشیاء به کار می‌رود [۱۲]. بیان اسکلت را می‌توان برای مدل کردن اشیاء صلب و همچنین اشیاء بندبند مورد استفاده قرار داد. (شکل ۲-۱ و).
راههای مختلفی برای بیان ویژگی‌های ظاهر شیء وجود دارد. برای دنبال کردن بیان شکل را می‌توان با بیان ظاهر ترکیب کرد [۱۳]. در ادامه بعضی از روشهای متداول برای بیان ظاهر که در الگوریتمهای دنبال کردن شیء به کار برده می‌شوند آورده شده است:

چگالی احتمال ظاهر شیء^{۱۲}: چگالی احتمال ویژگی‌های ظاهر شیء (رنگ، بافت^{۱۳}) را می‌توان برای ناحیه‌ای از تصویر که با مدلهای شکل (مثلاً ناحیه درونی یک بیضی یا یک کانتور) مشخص شده‌است، محاسبه کرد. روشهای تخمین چگالی احتمال ویژگی‌های ظاهر شیء را می‌توان به دو دسته روشهای

¹¹ Articulated shape models

¹² Object appearance

¹³ Texture

پارامتریک مثل گوسین [۱۴] یا تلفیقی از گوسین‌ها [۱۵] یا غیر پارامتریک مثل پنجره‌های پارزن^{۱۴} [۱۶] و هیستوگرام‌ها [۲] تقسیم نمود .

الگوها^{۱۵}: الگوها توسط اشکال ساده هندسی یا سایه نما تشکیل می‌شوند [۱۷]. یک مزیت الگو این است که هم اطلاعات مکانی و هم اطلاعات ظاهر را با خود دارند. الگوها ظاهر شیء را از یک نما در نظر می‌گیرند بنابراین آنها فقط برای مواردی که نمای مشاهده شده از شیء، در طول فرایند دنبال کردن، تغییر چندانی ندارد، مناسب هستند.

مدلهای فعال ظاهر^{۱۶}: مدلهای فعال ظاهر با مدل کردن همزمان شکل و ظاهر شیء تولید می‌شوند [۱۸]. در این روش در حالت کلی شکل شیء توسط مجموعه ای از علامتهای مشخصه^{۱۷} تعریف می‌شود . علامتهای مشخصه می‌توانند بر روی مرز شیء یا درون ناحیه شیء قرار گیرند. برای هر علامت مشخصه، یک بردار ظاهر^{۱۸} که ویژگی های رنگ، بافت یا دامنه گرادیان را در بر دارد ذخیره می‌شود.

مدلهای ظاهر چندنمایی^{۱۹}: این مدلها نماهای مختلفی از یک شیء را در نظر می‌گیرند. یک راهکار برای نمایش نماهای مختلف شیء ، تولید زیر فضایی از نماها می‌باشد. روشهای زیر فضا به طور مثال تحلیل مولفه

¹⁴ Parzen windows

¹⁵ Templates

¹⁶ Active appearance models

¹⁷ Landmarks

¹⁸ Appearance vector

¹⁹ Multiview appearance models

اصلی^{۲۰} و تحلیل مولفه مستقل^{۲۱} را هم برای بیان ظاهر و هم برای بیان شکل می توان به کار برد [۱۹ و ۲۰].

یک روش دیگر برای یادگیری نماهای مختلف از یک شیء، آموزش مجموعه‌ای از طبقه‌بندها برای مثال

بردارهای ماشین پشتیبان^{۲۲} [۲۱] یا شبکه‌های بیزین^{۲۳} [۲۲] می‌باشد.

بیان شیء معمولاً بر حسب حوزه کاربرد انتخاب می‌شود. مثلاً برای دنبال کردن اشیائی که در یک تصویر به

صورت بسیار کوچک ظاهر می‌شوند، معمولاً بیان نقطه مناسب است.

۲-۳ انتخاب ویژگی برای دنبال کردن شیء

انتخاب صحیح ویژگی‌ها، نقش کلیدی در طراحی یک الگوریتم دنبال کردن شیء ایفا می‌کند. در حالت کلی

در زمینه دنبال کردن شیء، مطلوب ترین خصوصیت یک ویژگی بصری، منحصر به فرد بودن آن است به

گونه‌ای که اشیاء را بتوان به راحتی در فضای ویژگی از هم تشخیص داد.

²⁰ Principal Component Analysis (PCA)

²¹ Independent Component Analysis (ICA)

²² Support vector machines

²³ Bayesian networks

انتخاب ویژگی رابطه تنگاتنگی با بیان شیء دارد. به طور مثال برای بیان ظاهر مبتنی بر هیستوگرام، رنگ را می‌توان به عنوان ویژگی به کار برد. در حالی که برای بیان مبتنی بر کانتور، لبه‌های شیء معمولاً به عنوان ویژگی به کار گرفته می‌شوند [۷]. بسیاری از الگوریتمهای دنبال کردن نیز ترکیبی از این ویژگی‌ها را به کار می‌گیرند.

در این قسمت تعدادی از ویژگی‌های بصری متداول را مطرح می‌کنیم:

رنگ: رنگ ظاهری یک شیء عمدتاً تحت تأثیر دو عامل فیزیکی توزیع طیف توان تابش کننده و ویژگیهای بازتاب کنندگی سطح شیء قرار دارد. در پردازش تصویر معمولاً فضای رنگ RGB برای بیان رنگ به کار می‌رود. لیکن فضای RGB یک فضای رنگ یک شکل از لحاظ ادراکی^{۲۴} نمی‌باشد به این معنی که اختلاف بین رنگها در فضای RGB متناظر با اختلافی که بین رنگها توسط انسان ادراک می‌شود نیست [۲۳] علاوه بر این مولفه های B و G و R شدیداً همبسته هستند. فضاهای رنگ $L^*u^*v^*$ و $L^*a^*b^*$ به گونه ای طراحی شده اند که "از لحاظ ادراکی یک شکل بودن" را تقریب بزند و فضای رنگ HSV تقریباً از لحاظ ادراکی یک شکل می‌باشد هر چند این فضاهای رنگی به نوبز حساسند [۲۴].

لبه‌ها: مرزهای شیء معمولاً تغییرات زیادی در شدت روشنایی تصویر ایجاد می‌کنند. آشکار سازی لبه برای تعیین این تغییرات به کار می‌رود. یک ویژگی مهم لبه این است که در مقایسه با مشخصه‌های رنگ در مقابل

²⁴ Perceptually uniform

تغییرات روشنایی حساسیت کمتری دارد. الگوریتمهایی که مرز شیء را دنبال می‌کنند معمولاً از لبه به

عنوان ویژگی استفاده می‌کنند. یکی از محبوب‌ترین روشهای آشکارسازی لبه، روش `canny` [۲۵] می‌باشد.

جریان نوری^{۲۵}: جریان نوری یک حوزه متراکم از بردارهای جابجایی می‌باشد که این بردارها انتقال هر

پیکسل را در یک ناحیه از یک فریم به فریم بعدی تعریف می‌کنند. جریان نوری با استفاده از قید روشنی^{۲۶}

محاسبه می‌شود. در قید روشنی فرض می‌شود که روشنی پیکسلها در فریمها متوالی ثابت است. استفاده از

جریان نوری به عنوان یک ویژگی در کاربرد های دنبال کردن شیء و بخش بندی مبتنی بر حرکت، متداول

است.

از تکنیکهای مشهور برای محاسبه جریان نوری متراکم می‌توان به [۲۶، ۲۷، ۲۸، ۲۹] اشاره کرد. همچنین

ارزیابی مفصل روشهای جریان نوری، در مرجع [۳۰] آمده است. به عنوان نمونه برای جریان نوری در

شکل ۲-۲ تصاویری مربوط به دو رشته ویدئویی و در شکل ۲-۳ جریانهای نوری مربوط به آن آورده شده است.

²⁵ Optical Flow

^{۲۶} Brightness constraint

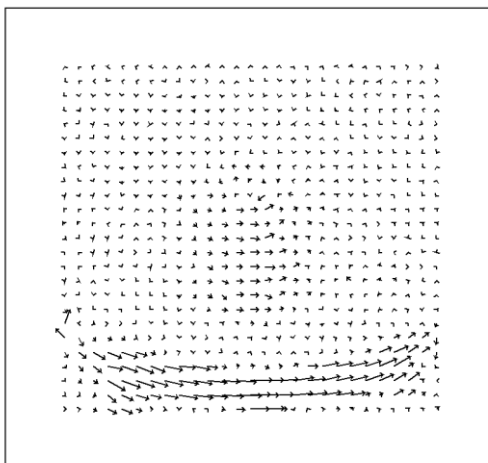


ب

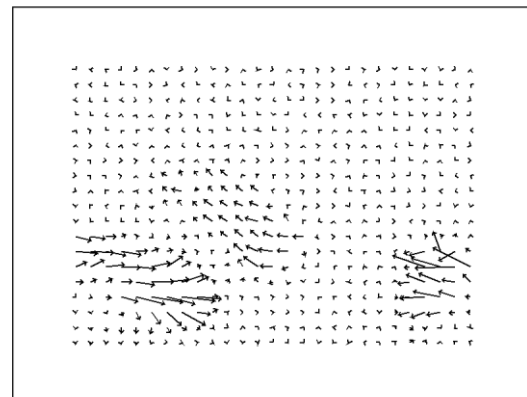


الف

شکل ۲-۲ تصویری از یک رشته ویدئویی الف) در این رشته چهار شیء متحرک داریم: ۱: ماشینی که در حال گردش کردن به سمت راست است. ۲: ماشینی در قسمت پایین سمت چپ که از چپ به راست حرکت می کند. ۳: خودروی ون در قسمت پایین سمت راست که از راست به چپ حرکت می کند. ۴: عابر در قسمت بالا و سمت چپ تصویر. ب) مکعبی بر روی یک صفحه دوار که برخلاف جهت عقربه های ساعت در حال چرخش است.



ب



الف

شکل ۳-۲ جریان های نوری مربوط به شکل ۲-۲

بافت^{۲۷}: بافت، به تغییرات روشنایی ساختاریافته گفته می شود مانند صفحه شطرنج و یا تصویر پارچه طرح

دار. در مقایسه با رنگ ، توصیف بافت نیاز به پردازش بیشتری دارد. مشابه ویژگی لبه ، توصیف گرهای بافت

نیز در مقایسه با رنگ در مقابل تغییرات روشنایی حساسیت کمتری دارند.

۲-۴ آشکار سازی شیء

در بعضی از کاربردهای مربوط به دنبال کردن شیء، هنگامی که برای اولین بار شیء وارد زاویه دید دوربین

می شود ، باید آشکارسازی شود و در بعضی از کاربردها نیز که چند شیء باید به طور همزمان دنبال شوند،

در هر فریم آشکارسازی صورت می گیرد 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 می باشند.

یک روش متداول برای آشکارسازی شیء متحرک در یک رشته ویدئویی استفاده از اطلاعات یک فریم می

باشد. در برخی از روشها نیز اطلاعات زمانی^{۲۸} که با در نظر گرفتن رشته ای از فریمها بدست می آید مورد

استفاده قرار داده می شود. این اطلاعات زمانی ، معمولاً به شکل تفاضل فریمها هستند که نواحی تغییر کننده

در فریمهای متوالی را برجسته می کند.

Texture²⁷

Temporal information²⁸

۲-۴-۱ آشکارسازی نقاط مهم

این روشها برای یافتن نقاط مهم که در توصیف بافت و انطباق تصاویر نقش مؤثری دارند، به کار می‌روند. از

جمله این آشکارسازهای توان به آشکارسازهای Harris [۳۱]، KLT [۳۲] و SIFT [۳۳] اشاره کرد.

در شکل ۲-۴ تصویر یک صحنه و نقاط آشکارسازی شده در آن به سه روش مختلف آورده شده است.



ج

ب

الف

شکل ۲-۴ نقاط مهم آشکارسازی شده با اعمال عملگرهای الف (SIFT ب) KLT ج) Harris

۲-۴-۲ تفریق پس زمینه

آشکار سازی شیء را می‌توان با تهیه مدلی برای پس زمینه و سپس یافتن انحرافها از این مدل برای هر

فریم ورودی، انجام داد. هر تغییر مهمی در ناحیه‌ای از تصویر نسبت به مدل پس زمینه، دلالت بر وجود یک

شیء متحرک دارد. پیکسلهایی که ناحیه تغییر کننده را تشکیل می‌دهند، برای پردازشهای بعدی علامت

گذاری می‌شوند. معمولاً یک الگوریتم بخش متصل^{۲۹} برای بدست آوردن نواحی متصل متناظر با اشیاء ،

اعمال می‌شود. در شکل ۲-۵ نمونه ای از خروجی آشکار سازی از طریق تفریق پس زمینه آورده شده است.



ج

ب

الف

شکل ۲-۵ آشکار سازی با استفاده از تفریق پس زمینه (الف) تصویر اصلی (ب) تصویر پس زمینه (ج) نتیجه تفریق پس زمینه

۲-۴-۳ بخش بندی^{۳۰} تصویر

هدف الگوریتمهای بخش بندی تصویر، قسمت بندی تصویر به ناحیه هایی است که از لحاظ ادراکی مشابهند

شکل (۲-۶). هر الگوریتم بخش بندی دو مسئله را مشخص می کند اول معیاری برای یک بخش خوب و دوم

روشی کارآمد برای قسمت بندی^[۳۴]. از جمله روشهای بخش بندی می توان به روشهای بخش بندی با

استفاده از روش انتقال میانگین^{۳۱} ، گراف کاتها^{۳۲} و کانتورهای فعال^{۳۳} اشاره کرد. شکل ۲-۶ نمونه از بخش

بندی تصویر با استفاده از روش انتقال میانگین را نشان می دهد [۱] .

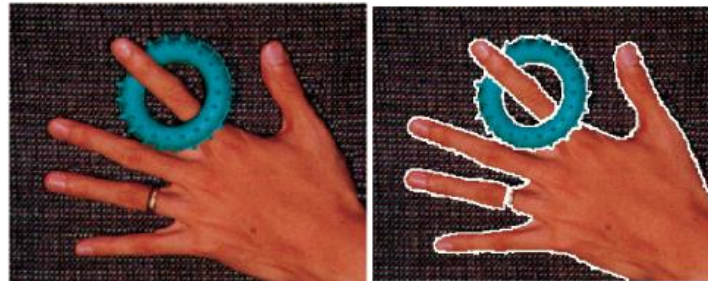
Connected component algorithm²⁹

Segmentation³⁰

Mean-Shift³¹

Graph-Cuts.³²

Active Contours³³



ب

الف

شکل ۲-۶ نمونه ای از بخش بندی تصویر با استفاده از روش انتقال میانگین (الف) تصویر اصلی (ب) تصویر بخش بندی شده

۲-۴-۴ یادگیری با سرپرست^{۳۴}

آشکارسازی 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^{۳۵} را از بین می برد. با استفاده از مجموعه ای از مثالهای محدود یادگیری، این روش قادر است نماهای نادیده شیء را بیاموزد. در زمینه آشکارسازی شیء مثالهای یادگیری ترکیب شده اند از زوجهایی از ویژگی های شیء و کلاس شیء اختصاص داده شده به آنها که هر دوی این کمیت ها به صورت دستی تعریف می شوند [۷]. 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 شدند، می توان یادگیری ظاهرهای

³⁴ Supervised Learning

³⁵ Templates

مختلف از یک شیء را با استفاده از یک روش یادگیری با سرپرست انجام داد. شبکه‌های عصبی^{۳۶} [۳۵] بوستینگ تطبیقی^{۳۷} [۳۶] درختهای تصمیم^{۳۸} [۳۷] و ماشین بردار پشتیبان^{۳۹} [۳۸] از جمله روشهای کارآمد در این زمینه می باشند. این روشهای یادگیری، یک ابر رویه^{۴۰} که یک کلاس شیء را از کلاسهای شیء دیگر جدا می کند را محاسبه می کنند.

۲-۵ دنبال کردن شیء

هدف دنبال کننده شیء در یک رشته ویدئویی، تعیین مسیر حرکت شیء در فریمهای متوالی آن رشته می باشد. دنبال کننده شیء ممکن است ، کل ناحیه ای را که در هر لحظه زمانی توسط شیء اشغال می شود را نیز بدست دهد. آشکارسازی شیء و برقراری تناظر بین اشیاء در فریمها را می توان به طور جداگانه یا در تلفیق با یکدیگر انجام داد. در مورد اول، نواحی ممکن برای شیء در هر فریم توسط الگوریتم آشکارسازی شیء تعیین می شود و سپس دنبال کننده تناظر اشیاء را در فریمها تعیین می کند. در مورد دوم با به روز کردن مکرر موقعیت شیء و اطلاعات ناحیه که از فریمهای قبلی بدست آمده تناظر و ناحیه شیء با همدیگر تخمین زده می شوند [۷].

Neural networks³⁶

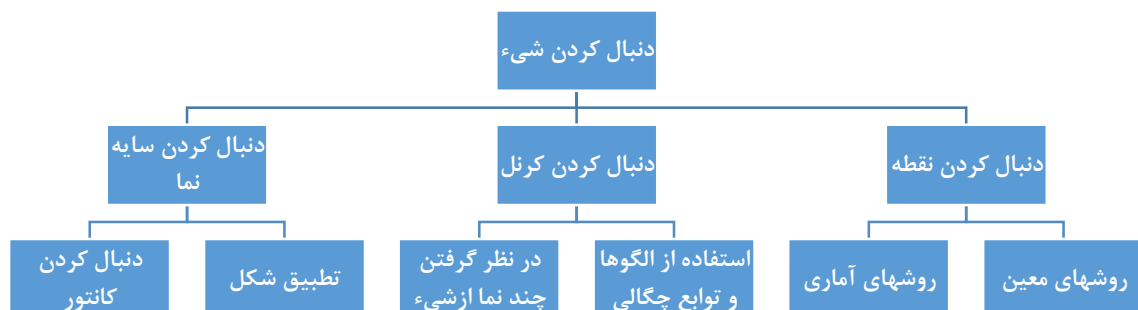
Adaptive boosting³⁷

Decision trees³⁸

Support vector machine³⁹

Hypersurface⁴⁰

مدلی که برای بیان شکل شیء انتخاب شده است نوع حرکت و یا تغییر شکل شیء را محدود می‌سازد. برای مثال اگر شیء با یک نقطه بیان شده باشد فقط می‌توان از مدل انتقال استفاده کرد و هنگامی که یک بیان شکل هندسی مانند بیضی برای شیء استفاده شود علاوه بر مدل انتقال مدلهای حرکت پارامتریک مانند تبدیلات پروژکتیو^{۴۱} یا افاین نیز قابل استفاده هستند، این بیانها می‌توانند حرکت اجسام صلب را در صحنه تقریب بزنند. برای بیان اشیاء غیرصلب می‌توان از سایه نما یا کانتور استفاده کرد و مدلهای پارامتریک و غیر پارامتریک را برای تعیین حرکت آنها به کار برد. در شکل ۲-۷ دسته بندی روشهای دنبال کردن صورت گرفته است.

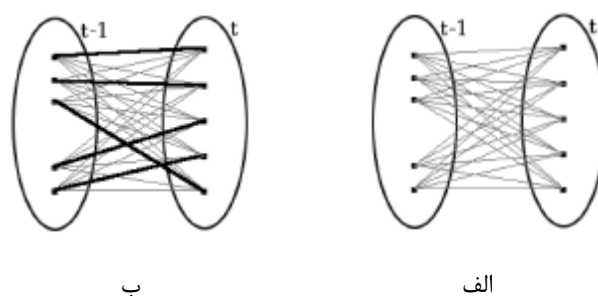


شکل ۲-۷ دسته بندی روشهای دنبال کردن شیء

۲-۵-۱ دنبال کردن نقطه

در روشهای دنبال کردن نقطه، شیء آشکارسازی شده در فریمهای متوالی با نقاط بیان می‌شود و دنبال کننده باید تعیین کند که هر نقطه آشکارسازی شده با کدام نقطه در فریم قبلی یا سایر فریمها متناظر است. این راهکار نیاز به یک مکانیزم خارجی برای آشکارسازی شیء در هر فریم دارد. دنبال کردن نقطه در واقع تعیین تناظر بین نقاط استخراج شده، در فریمهای رشته ویدئویی می باشد. تعیین تناظر نقاط در دو یا چند تصویر یک مسئله پیچیده است به ویژه در مواردی که احتمال ورود یا خروج اشیاء یا آشکارسازی اشتباه و یا همپوشانی اشیاء متحرک وجود دارد.

به طور کلی روشهای تعیین تناظر نقطه را می‌توان به دو دسته روشهای معین^{۴۲} و روشهای آماری تقسیم کرد. روش معین از راهنماهای حرکت کیفی^{۴۳} [۸] برای محدود کردن مسئله تناظر استفاده می‌کند. در حالی که روشهای آماری صراحتاً مشاهدات مربوط به شیء را در نظر می‌گیرند و عدم قطعیت را برای برقراری تناظر وارد محاسبات می‌کنند.



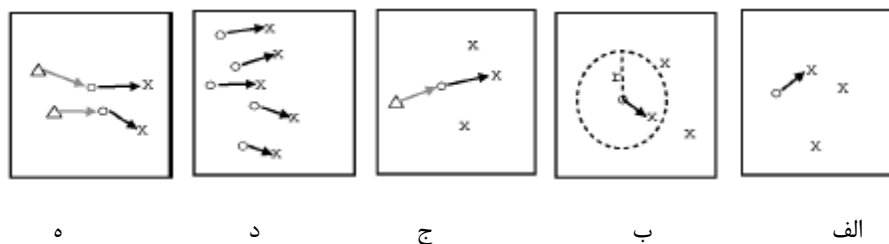
شکل ۲-۸ برقراری تناظر بین نقاط (الف) تمامی تناظرات ممکن بین نقاط در دوفریم متوالی (ب) تناظر یک به یک

Deterministic Method ⁴²

Qualitative motion heuristics ⁴³

۲-۵-۱-۱ روشهای معین برای تعیین تناظر

این روشها هزینه‌ای برای نظیر کردن هر شیء در فریم $t-1$ به یک شیء در فریم t با استفاده از قیدهای حرکت، تعریف می‌کنند. می‌نیمم سازی این تابع هزینه به شکل یک مسئله بهینه سازی ترکیبی فرموله می‌شود. یک راه حل که مشتمل بر تناظر نقطه به نقطه (شکل ۲-۸ب) در میان تمامی تناظرهای ممکن (شکل ۲-۸الف) است را می‌توان با روشهای تخصیص بهینه مانند الگوریتم هانگاریان^{۴۴} [۳۹] یا روشهای جستجوی آزمند^{۴۵}، بدست آورد. هزینه تناظر معمولاً با استفاده از ترکیبی از قیدهای زیر تعریف می‌شود.



شکل ۲-۹ قیود حرکتی مختلف (الف) مجاورت (ب) حداکثر سرعت (ج) اندک بودن تغییرات سرعت (د) حرکت مشترک (ه) صلب بودن

مجاورت^{۴۶}: در قید مجاورت فرض می‌شود که موقعیت شیء از یک فریم به فریم دیگر تغییر چشمگیری نخواهد داشت (شکل ۲-۹الف).

⁴⁴ Hungarian algorithm

⁴⁵ Greedy

⁴⁶ Proximity

حداکثر سرعت: یک حد بالا برای سرعت شیء در نظر می گیرد و تناظرهای ممکن را به یک همسایگی

دایره شکل در اطراف شیء محدود می کند (شکل ۲-۹ب)

اندک بودن تغییرات سرعت (حرکت هموار^{۴۷}): در این روش فرض می شود که جهت و سرعت شیء

تغییرات زیادی نخواهد داشت. (شکل ۲-۹ج)

حرکت مشترک: فرض می کند که سرعت اشیاء در یک همسایگی باید مشابه باشند (شکل ۲-۹د) این قید

برای اشیاء مناسب است که با چندین نقطه بیان می شوند.

صلب بودن: در این قید فرض می شود که اشیاء در دنیای سه بعدی صلب هستند بنابراین فاصله بین هر دو

نقطه بر روی شیء واقعی بدون تغییر باقی می ماند (شکل ۲-۹هـ) [۷].

یک شکلی مجاورتی^{۴۸}: این قید ترکیبی از قیود مجاورت و تغییرات اندک سرعت می باشد.

در اینجا لازم به یادآوری است که این قیود مختص به روشهای معین نیستند و می توانند در دنبال کردن

نقطه با استفاده از روشهای آماری نیز مورد استفاده قرار گیرند.

۲-۱-۵-۲ روشهای آماری برای تعیین تناظر

⁴⁷ Smooth

⁴⁸ Proximal uniformity

سنجش‌هایی که از حسگرهای ویدئویی بدست آمده‌اند همیشه دارای نویز می‌باشند. علاوه بر این حرکت شیء ممکن است دستخوش اغتشاشات تصادفی قرار گیرد مانند یک وسیله نقلیه که در حال مانور دادن است . روشهای تناظر آماری این مشکلات دنبال کردن را به این شکل حل می‌کنند که سنجش‌ها را در نظر می‌گیرند و عدم قطعیت مدل را در طول تخمین وارد محاسبات می‌کنند. روشهای تناظر آماری از راهکار فضای حالت^{۴۹} برای مدل کردن ویژگیهایی از شیء مانند موقعیت، سرعت و شتاب استفاده می‌کنند. سنجشها معمولاً شامل موقعیت شیء در تصویر که توسط مکانیزم آشکارسازی بدست می‌آید، می‌باشند.

برای تخمین حالت یک شیء منفرد می‌توان از فیلتر ذره ای^{۵۰} استفاده کرد[۴۰] و در صورت در نظر گرفتن برخی فرضهای محدود کننده خاص، تخمین حالت بهینه با فیلتر کالمن^{۵۱} صورت می‌گیرد . در مرجع [۴۱] فیلتر کالمن برای دنبال کردن نقاط در یک تصویر نویزی به کار گرفته شده است . در مرجع [۴۲] از فیلتر کالمن توسعه یافته برای تخمین مسیر سه بعدی یک جسم از حرکت دوبعدی، استفاده شده است. فیلتر کالمن و فیلتر ذره ای در هر لحظه یک سنجش واحد را فرض می‌کنند زیرا حالت یک شیء واحد تخمین زده می‌شود. دنبال کردن چند شیء نیاز به راه حلی تلفیقی برای مسائل تخصیص داده و تخمین حالت دارد. اگرچه در اینجا روشهای فیلتر کالمن و فیلتر ذره ای را در بخش دنبال کردن نقطه مطرح کردیم ، این روشها در دنبال کردن کرنل وسایه نما نیز می‌توانند مورد استفاده قرار گیرند.

State space ⁴⁹

Particle filter ⁵⁰

Kalman Filter ⁵¹

یک مسئله مهم در زمینه دنبال کننده‌های نقطه برخورد با مسائل مشاهدات نویزی و مشاهدات از دست رفته می‌باشد. برای برخورد با این مسائل دنبال کننده‌های نقطه که از روشهای معین برای تعیین تناظر استفاده می‌کنند اغلب ترکیبی از قیود مبتنی بر حرکت مانند حرکت مشترک [۸] و یک شکلی مجاورتی [۴۳] را به کار می‌گیرند.

روشهای آماری دنبال کردن نقطه، با لحاظ کردن عدم قطعیت مدل، صراحتاً به مسئله نویز می‌پردازند. در بعضی موارد این فرض که سنجشها به صورت نرمال حول موقعیت پیش‌بینی شده شان توزیع می‌شوند، ممکن است لحاظ نشود. علاوه بر این در بسیاری از موارد پارامترهای نویز را نمی‌دانیم.

دنبال کننده‌های نقطه برای دنبال کردن اشیاء خیلی کوچک که می‌توان آنها را با یک نقطه بیان کرد (بیان تک نقطه‌ای) مناسب هستند. برای دنبال کردن اشیاء بزرگتر چند نقطه مورد نیاز است. در زمینه دنبال کردن شیء با استفاده از چند نقطه برای تمیز دادن بین چند شیء و یا تمیز دادن بین اشیاء و پس زمینه می‌توان از روشهای خوشه‌بندی اتوماتیک برای تعیین نقاطی که متعلق به یک شیء هستند استفاده کرد. روشهای بخش‌بندی و خوشه‌بندی مبتنی بر حرکت [۴۴ و ۲۸ و ۴۵] برای ساده‌سازی مسئله بخش‌بندی معمولاً فرض می‌کنند نقاطی که دنبال می‌شوند بر روی یک بدنه صلب قرار دارند.

۲-۵-۲ دنبال کردن کرنل^{۵۲}

کرنل به شکل و ظاهر شیء اشاره دارد. برای مثال کرنل می‌تواند الگوی^{۵۳} مستطیلی یا یک شکل بیضوی به همراه هیستوگرام رنگ مربوط به ناحیه داخل آن باشد. دنبال کردن شیء با محاسبه حرکت کرنل در فریمهای متوالی، صورت می‌گیرد. در روشهای دنبال کردن کرنل حرکت شیء معمولاً به شکل تبدیلات هندسی پارامتریک مانند انتقال، چرخش و افاین در نظر گرفته می‌شود.

دنبال کردن کرنل نوعاً با محاسبه حرکت شیء از یک فریم به فریم بعدی انجام می‌پذیرد. این الگوریتمها بر حسب بیان ظاهر بکار رفته، تعداد اشیائی که دنبال می‌شوند و روشی که برای تخمین حرکت شیء به کار می‌رود از یکدیگر متمایز می‌شوند.

در این تحقیق روشهای دنبال کردن کرنل را بر حسب روشی که برای بیان ظاهر شیء به کار می‌گیرند به دو زیر گروه « روشهایی که از مدلهای ظاهر مبتنی بر چگالی و الگوها^{۵۴} استفاده می‌کنند» و « روشهایی که از مدلهای ظاهر چند نمایی^{۵۵} استفاده می‌کنند» تقسیم می‌کنیم.

۲-۵-۱-۲ دنبال کردن با استفاده از الگوها و مدلهای مبتنی بر تابع چگالی ویژگیهای ظاهر شیء

الگوها و مدلهای ظاهر مبتنی بر چگالی ویژگی های تصویر شیء به خاطر اینکه نسبتاً ساده اند و هزینه محاسباتی کمی دارند، به طور گسترده‌ای مورد استفاده قرار گرفته‌اند. یکی از روشهای متداول در این بخش

Template⁵³

Templates and density-based appearance models⁵⁴

Multiview appearance models⁵⁵

برای دنبال کردن اشیاء منفرد، تطبیق الگو^{۵۶} می‌باشد. تطبیق الگو یک روش brute force جستجوی تصویر

I_w برای ناحیه‌ای مشابه الگوی شیء (O_t) که در فریم قبلی تعریف شده، می‌باشد. موقعیت الگو در فریم

جاری با محاسبه یک معیار مشابهت^{۵۷} محاسبه می‌شود. برای مثال این معیار مشابهت می‌تواند همبستگی

تقاطع^{۵۸} باشد:

$$\arg \max_{dx, dy} \frac{\sum_x \sum_y (O_t(x, y) \times I_w(x + dx, y + dy))}{\sqrt{\sum_x \sum_y O_t^2(x, y)}} \quad (1-2)$$

که در آن (dx, dy) موقعیت الگوی کاندید را تعیین می‌کند.

معمولاً شدت روشنایی تصویر یا ویژگی‌های رنگ برای توصیف الگو به کار می‌روند. از آنجا که شدت روشنایی

تصویر به تغییرات روشنایی خیلی حساس است می‌توان از گرادیانهای شدت روشنایی تصویر [۴۶] نیز به

عنوان ویژگی استفاده کرد. یک محدودیت برای تطبیق الگو هزینه محاسباتی بالا به خاطر جستجوی brute

force می‌باشد. برای کاهش هزینه محاسباتی، جستجو معمولاً به مجاورت موقعیت قبلی شیء محدود

می‌شود. همچنین الگوریتمهای کارآمدتری برای تطبیق الگو پیشنهاد شده‌اند [۴۷].

برای دنبال کردن شیء، به جای الگوها بیانهای دیگری نیز قابل استفاده‌اند. برای مثال هیستوگرامهای

رنگ یا مدل‌های ترکیبی را می‌توان با استفاده از ظاهر پیکسل‌های درون نواحی بیضوی یا مستطیلی، محاسبه

Template matching ⁵⁶

Similarity measure ⁵⁷

Cross correlation ⁵⁸

کرد. به طور مثال در مرجع [۱۷] مدل شیء با یافتن رنگ میانگین پیکسلهای داخل ناحیه شیء که با مستطیل مشخص می‌شود، تعیین می‌گردد. در این روش برای کاهش پیچیدگی محاسباتی جستجو برای شیء فقط به هشت موقعیت همسایه محدود می‌شود. مشابهت بین مدل شیء (M) و مدل ناحیه کاندید (H) با محاسبه نسبت بین میانگینهای رنگ بدست آمده از H, M اندازه‌گیری می‌گردد. موقعیتی که بالاترین نسبت را به خود اختصاص دهد به عنوان موقعیت جاری شیء انتخاب می‌شود.

۲-۵-۲-۲ دنبال کردن با استفاده از مدل‌های ظاهر چند نمایی^{۵۹}

در روشهای مرور شده، مدل‌های ظاهر شیء که هیستوگرامها، الگوها و غیره هستند معمولاً به صورت برخط^{۶۰} تولید می‌شوند. بنابراین این مدلها اطلاعاتی از شیء را که از آخرین مشاهدات جمع‌آوری شده‌اند را ارائه می‌دهند. اشیاء ممکن است از نماهای متفاوت، ظاهری متفاوت داشته باشند. و اگر نمای شیء در طول دنبال کردن زیادی کند ممکن است مدل ظاهر دیگر معتبر نباشد و مسیر شیء ممکن است گم شود. برای فائق آمدن بر این مشکل می‌توان نماهای مختلفی از شیء را به صورت برون خط^{۶۱} یادگیری کرد و برای دنبال کردن به کار برد.

یک مقایسه کیفی بین دنبال کننده‌های کرنل می‌تواند بر پایه موارد زیر صورت پذیرد: تعداد اشیائی که دنبال کننده می‌تواند دنبال کند، توانایی یا عدم توانایی در برخورد با مسئله مخفی شدن، نیاز به آموزش و

⁵⁹ Multiview Appearance Models

⁶⁰ Online

⁶¹ Offline

نوع مدل حرکت. حرکت شیء را می‌توان با ماکزیمم سازی مشابهت ظاهر شیء بین فریم جاری و فریم قبلی ، تخمین زد. فرایند تخمین می‌تواند به شکل یک جستجوی brute force یا با استفاده از فرایند ماکزیمم سازی (مینیمم سازی) مبتنی بر افزایش (کاهش) گرادیان باشد. دنبال کننده‌های هدف مبتنی بر راهکار افزایش یا کاهش گرادیان نیاز دارند که حداقل بخشی از شیء درون شکل انتخاب شده که موقعیتش توسط موقعیت شیء پیشین تعریف می‌شود، قابل رویت باشد [۷]. برای رفع چنین نیازهایی ، یک راهکار ممکن استفاده از فیلتر کالمن یا فیلتر ذره ای است که موقعیت شیء در فریم بعدی را پیش بینی می‌کند.

یک محدودیت بیان شیء توسط اشکال هندسی پایه این است که قسمتهایی از شیء ممکن است بیرون از شکل تعیین شده قرار گیرد و یا قسمتهایی از پس زمینه داخل شکل بیافتد . این پدیده را می‌توان هم در اشیاء صلب هنگامی که موقعیت شیء در مقابل دوربین تغییر می‌کند و هم در اشیاء غیرصلب، وقتی حرکت اجزاء آن نسبت به یکدیگر باعث ایجاد تغییر در ظاهر شیء می‌شود مشاهده نمود. در چنین مواردی تخمین حرکت شیء که با ماکزیمم‌سازی مشابهت مدل بدست آمده، ممکن است صحیح نباشد. برای فائق آمدن بر این مشکل، یک راهکار این است که کرنل را طوری تعیین کنیم که به جای احاطه کل شیء، درون شیء قرار گیرد .

۲-۵-۳ دنبال کردن سایه نما^{۶۲}

⁶² Silhouette Tracking

اشیاء ممکن است ، اشکال پیچیده‌ای داشته باشند برای مثال دستها، سر و شانه‌ها که نمی‌توان آنها را با اشکال هندسی ساده توصیف کرد. روشهای مبتنی بر سایه نما توصیف دقیقی از شکل را برای این اشیاء مهیا می‌کنند و در مواردی که نیاز به دنبال کردن کل ناحیه شیء باشد می‌توان از این روشها استفاده کرد. در شکل ۱۰-۲ نمونه ای از یک تصویر و سایه نمای مربوط به شیء در آن تصویر، آورده شده است.

در این روشها دنبال کردن با تخمین ناحیه شیء در هر فریم صورت می‌گیرد . روشهای دنبال کردن سایه‌نما از اطلاعاتی که در داخل ناحیه شیء نهفته است استفاده می‌کنند. این اطلاعات می‌توانند به فرم چگالی ویژگیهای ظاهر شیء و مدل‌های شکل که معمولاً به صورت نگاشتهای لبه هستند باشند [۷].



شکل ۱۰-۲ (الف) شیء مورد نظر فردی است که جسمی را حمل می‌کند. (ب) سایه نمای مربوط به شیء

هدف یک دنبال کننده شیء مبتنی بر سایه‌نما ، تعیین ناحیه شیء در هر فریم می‌باشد. این امر توسط مدل شیء که با استفاده از فریمهای قبلی تولید شده تحقق می‌یابد. این مدلها می‌توانند با کمک هیستوگرام رنگ، لبه‌های شیء یا کانتور شیء ایجاد شوند. ما روشهایی که دنبال کننده های سایه‌نما به کار می‌گیرند را

به دو دسته تطبیق شکل و دنبال کردن کانتور تقسیم می‌کنیم. روشهای تطبیق شکل به جستجوی سایه‌نمای شیء در فریم جاری می‌پردازند. از سوی دیگر الگوریتمهای دنبال کننده کانتور توسط مدل‌های فضای حالت یا مینم‌سازی مستقیم یک تابع انرژی، کانتور تغییر یافته را در فریم جاری را بدست می‌آورند.

۲-۵-۳-۱ تطبیق شکل

تطبیق شکل را می‌توان مشابه دنبال کردن بر پایه تطبیق الگو انجام داد. به این شکل که سایه‌نما و مدل تخصیص داده شده به آن در فریم جاری مورد جستجو قرار گیرند. جستجو با محاسبه مشابهت ناحیه کاندید با مدل شیء صورت می‌گیرد. در این روش فرض می‌شود که سایه‌نما از فریم جاری به فریم بعدی فقط تحت تاثیر تبدیل هندسی انتقال قرار گرفته است بنابراین حرکت اشیاء غیر صلب مدل نمی‌شود. مدل شیء که معمولاً به شکل نگاشت لبه است برای حل مسئله تغییرات ظاهر، در هر فریم بعد از تعیین موقعیت شیء به روز می‌شود. این به روز رسانی برای فائق آمدن بر مشکلات مربوط به تغییرات شرایط نوری و زاویه دید و همچنین حرکت اجسام غیر صلب مورد نیاز است.

یک راهکار دیگر برای تطبیق شکلها یافتن سایه‌نماهای متناظر که در دو فریم متوالی آشکار سازی شده‌اند می‌باشد. برقرار کردن تناظر سایه نما یا به بیان خلاصه تطبیق سایه‌نما را می‌توان مشابه برقراری تناظر بین نقطه‌ها در نظر گرفت. هر چند تفاوت اصلی بین تطبیق سایه‌نما و تطبیق نقطه، بیانهای شیء و مدل‌های

شیء است که به کار رفته است . به ویژه اینکه تطبیق سایه‌نما به جای استفاده از نقاط، ناحیه کامل شیء را بکار می‌گیرد. علاوه بر این تطبیق سایه‌نما از ویژگی‌های ظاهر شیء استفاده می‌کند در حالی که در تطبیق نقطه تنها ویژگی‌های مبتنی بر موقعیت و حرکت به کار گرفته می‌شوند [۷]. آشکارسازی سایه‌نما معمولاً با استفاده از روش تفریق پس زمینه انجام می‌گیرد. هنگامی که سایه‌نمای شیء استخراج شد برای برقرار کردن تناظر بین سایه نماها می‌توان با استفاده از یک معیار مشابهت، شباهت بین سایه نماهای آشکارسازی شده در فریم جاری و سایه نماهای مربوط به فریم قبلی را تعیین کرد. مدل‌های شیء معمولاً به شکل توابع چگالی (هیستوگرامهای لبه یا رنگ) ، مرز سایه‌نما (کانتور شیء باز یا بسته) ، لبه‌های شیء یا ترکیبی از این مدل‌ها می‌باشد.

به جای جستجو برای تعیین سایه نماهای متناظر می‌توان دنبال کردن سایه‌نما را با محاسبه بردار جریان حرکت^{۶۳} برای هر پیکسل داخل سایه‌نما انجام داد. به این گونه بردار حرکتی که درون کل سایه‌نما غالب است مسیر حرکت سایه‌نما را تعیین می‌کند.

۲-۵-۳-۲ دنبال کردن کانتور

در روش‌های دنبال کردن کانتور در هر فریم یک کانتور اولیه در نظر گرفته می‌شود و در چند مرحله تغییراتی در این کانتور صورت می‌گیرد تا به بهترین مطابقت با شیء برسد. برای استفاده از این روشها باید

⁶³ Flow vector

بخشی از شیء در فریم جاری با ناحیه شیء در فریم قبلی همپوشانی داشته باشد. دنبال کردن مبتنی بر تغییر کانتور را می توان با استفاده از دو راهکار متفاوت انجام داد. اولین راهکار استفاده از مدل های فضای حالت برای مدل کردن حرکت و شکل کانتور می باشد. در دومین راهکار تغییر کانتور مستقیماً با مینیمم سازی انرژی کانتور انجام می شود. مینیمم سازی انرژی کانتور با استفاده از تکنیک های مینیمم سازی انرژی مانند کاهش گرادیان، صورت می گیرد.

دنبال کردن کانتور با استفاده از مدل های فضای حالت

در این روش ها حالت شیء بر حسب پارامترهای حرکت و شکل کانتور تعریف می شود و حالت در هر لحظه زمانی به گونه ای به روز می شود که احتمال پسین^{۶۴} کانتور ماکزیمم گردد. احتمال پسین به حالت پیشین و درست نمایی^{۶۵} فعلی بستگی دارد.

دنبال کردن کانتور با مینیمم سازی مستقیم تابع انرژی کانتور

بین روش های بخش بندی تصویر و روش های دنبال کردن با استفاده از تغییر دادن یک کانتور اولیه شباهتهایی وجود دارد. هم روش های بخش بندی و هم روش های دنبال کردن، تابع انرژی را با روش های آزمند^{۶۶} یا با نزول

⁶⁴ Posteriori probability

⁶⁵ Likelihood

⁶⁶ Greedy

گرادین مینیمم می کنند. انرژی کانتور بر حسب اطلاعات زمانی^{۶۷} به شکل گرادین زمانی (جریان نوری^{۶۸})

[۵۰ و ۴۹ و ۴۸] یا آمارهای ظاهر^{۶۹} که از نواحی پس زمینه و شیء تولید شده اند، [۵۱ و ۱۰] تعریف می شود.

کارهای گسترده ای که برای محاسبه جریان نوری صورت گرفته است، باعث ایجاد انگیزه برای دنبال کردن

کانتور با استفاده از گرادینهای تصویر زمانی شده است. قید جریان نوری از قید ثابت بودن روشنی گرفته شده

است:

$$I^{t+1}(x, y) - I^t(x - u, y - v) = 0 \quad (2-2)$$

که در آن I تابع شدت روشنایی تصویر، t زمان و (u, v) بردار جریان در جهتهای x, y می باشد.

جایگزینی برای جریان نوری به کارگیری مسئله ثابت بون آمارهای محاسبه شده بیرون و داخل ناحیه شیء

از یک فریم به فریم دیگر می باشد. در این راهکار باید موقعیت اولیه کانتور در فریم جاری به صورت موقعیت

کانتور در فریم قبلی در نظر گرفته شود.

موارد زیر از فاکتورهای مهم برای تفکیک کردن بین دنبال کننده های سایه نما می باشند: ویژگیهای استفاده

شده، چگونگی برخورد با مسئله مخفی شدن، نیاز یا عدم نیاز به آموزش و علاوه بر اینها بعضی از الگوریتمها

تنها اطلاعات مرز سایه نما را برای دنبال کردن به کار می گیرند در حالی که برخی، از اطلاعات تصویر ناحیه

⁶⁷ Temporal information

⁶⁸ Optical flow

⁶⁹ Appearance statistics

داخل سایه‌نما استفاده می‌کنند. راهکارهای مبتنی بر ناحیه عموماً در مقابل نویز انعطاف‌پذیری بیشتری دارند.

مهم‌ترین مزیت روشهای دنبال کردن مبتنی بر سایه‌نما، انعطاف‌پذیری آنها در دنبال کردن اشکال مختلف شیء می‌باشد. سایه‌نماها به صورت مختلفی قابل بیان هستند. متداول‌ترین بیان سایه‌نما به شکل تابع نشانگر باینری است که ناحیه شیء را با یک و نواحی غیر شیء را با صفر علامت‌گذاری می‌کند.

برای روشهای مبتنی بر کانتور بیان سایه‌نما به صورت صریح یا ضمنی می‌باشد. در بیان صریح مرز سایه‌نما با استفاده از مجموعه‌ای از نقاط کنترلی^{۷۰} تعریف می‌شود. بیان ضمنی، سایه‌نما را با استفاده از تابعی که بر روی یک شبکه تعریف شده، توصیف می‌کند. متداول‌ترین بیان کانتور به صورت ضمنی، بیان مجموعه‌های سطح^{۷۱} می‌باشد. ظاهر شیء معمولاً با توابع چگالی پارامتریک یا غیر پارامتریک مانند تلفیق گوسین‌ها یا هیستوگرامها، مدل می‌شود. شکل شیء را می‌توان به صورت زیر فضای کانتور مدل کرد. یک زیرفضا از مجموعه‌ای از کانتورهای ممکن شیء که از حالت‌های مختلف شیء بدست آمده‌اند تولید می‌شود [۵۲]. استفاده از بیانهای شکل مبتنی بر ظاهر، در روشهایی که جستجوی سایه‌نما را به صورت brute force انجام می‌دهند متداول است. برای بیانهای شکل مبتنی بر لبه، فاصله هاسدروف^{۷۲} معیاری است که گسترده‌ترین

Control points⁷⁰

⁷¹ Level Sets

⁷² Hausdorff distance

استفاده را دارد. لیکن معیار هاسدروف به نوبت حساس می باشد بنابراین به جای استفاده از بیشترین فاصله

محققین استفاده از میانگین فواصل را مد نظر قرار می دهند [۵۳].

مسئله مهم دیگر در زمینه روشهای دنبال کردن سایه نما چگونگی برخورد با مشکل مخفی شدن^{۷۳} شیء

مورد نظر توسط اجزاء دیگر می باشد. معمولاً روشها ارائه شده در مقالات مسئله مخفی شدن را صراحتاً

مورد بررسی قرار نمی دهند.

یکی از راهکارهای متداول فرض سرعت یا شتاب ثابت است که در آن ، موقعیت شیء تخمین زده می شود.

روشهای محدودی هستند که با اعمال قیود شکلی ، صراحتاً با مسئله مخفی شدن برخورد می کنند [۵۴ و ۱۰]

. یک مسئله دیگر در ارتباط با دنبال کننده های سایه نما توانایی آنها در برخورد با ادغام اشیاء و یا جدا شدن

بخشی از شیء می باشد. برای مثال در دنبال کردن شخصی که کیفی را حمل می کند ، هنگامی که شخص ،

کیف را رها می کند بخشی از کانتور شخصی بر روی شیء رها شده قرار می گیرد (جدا شدن ناحیه) این

تغییرات توپولوژی به صورت ادغام یا جدا شدن ناحیه را می توان با بیانهای ضمنی کانتور حل کرد.

۲-۶ مسائل مرتبط

۲-۶-۱ حل مسئله مخفی شدن شیء

⁷³ Occlusion

موارد مخفی شدن شیء را می‌توان در سه گروه طبقه‌بندی کرد: خود مخفی شدن ، مخفی شدن بین اشیاء و مخفی شدن توسط پس زمینه. خود مخفی شدن وقتی اتفاق می‌افتد که قسمتی از شیء جلو قسمت دیگر را بگیرد. این شرایط در طول دنبال کردن اشیاء بندبند زیاد اتفاق می‌افتد. مخفی شدن بین اشیاء هنگامی اتفاق می‌افتد که دو شیئی که دنبال می‌شوند جلوی یکدیگر را بگیرند. مخفی شدن با پس زمینه هنگامی اتفاق می‌افتد که قسمتی از پس زمینه جلوی شیئی که دنبال می‌شود را سد کند.

عموماً برای مخفی شدن بین اشیاء، دنبال کننده‌های چند شیء مانند [۵۴] و [۱۶] اطلاعات مربوط به موقعیت و ظاهر دو شیء مخفی شده و مخفی کننده (شیئی که در تصویر در جلوی شیء موردنظر قرار گرفته و باعث مخفی شدن آن شده است) را به کار می‌گیرند. آشکارسازی مخفی شدن جزئی یک شیء توسط صحنه مشکل است زیرا متمایز کردن آن از شیئی که شکلش در حال تغییر است کاری مشکل می‌باشد. یک راهکار برای برخورد با مسئله مخفی شدن کامل در طول دنبال کردن این است که حرکت شیء را با مدل‌های دینامیکی خطی یا غیر خطی مدل کنیم و هنگامی که مخفی شدن رخ می‌دهد، تا زمانی که شیء دوباره ظاهر می‌شود به پیش‌بینی موقعیت شیء ادامه دهیم. برای مثال یک مدل سرعت خطی در [۵۵] به کار گرفته شده است و یک کالمن فیلتر برای تخمین موقعیت و حرکت شیء استفاده شده است. همچنین در مرجع [۵۶] به کمک یک مدل دینامیک غیر خطی و یک فیلتر ذره ای حرکت شیء تخمین زده می‌شود.

احتمال رخداد مخفی شدن را می‌توان با انتخاب مناسب موقعیت دوربین کاهش داد. برای مثال اگر دوربینها بر روی یک وسیله نقلیه هوایی قرار گیرند و یک نما که از بالا گرفته شده است از صحنه در دسترس باشد مخفی شدن بین اشیاء بر روی زمین رخ نخواهد داد. هر چند در صورت استفاده از دوربین‌های با نمای اریب ممکن است با مخفی شدن چند شیء مواجه شویم و نیاز به مکانیزمی برای حل مسئله مخفی شدن باشد. برای حل مسئله مخفی شدن در طول دنبال کردن می‌توان از چند دوربین که یک صحنه را در زاویه دید خود دارند استفاده کرد [۵۷و۵۸].

۲-۶-۲ دنبال کردن با استفاده از چند دوربین

استفاده از چند دوربین برای دنبال کردن به دو دلیل می‌تواند مفید باشد . دلیل اول استفاده از اطلاعات عمق برای دنبال کردن و حل مسئله مخفی شدن است و دلیل دوم برای وسیعتر کردن ناحیه تحت نظر می‌باشد، زیرا یک دوربین به خاطر حوزه دید محدود سنسورها نمی‌تواند نواحی وسیعی را تصویر برداری کند. یک مسئله مهم در استفاده از چند دوربین ارتباط بین نماهای دوربینهای مختلف است که می‌توان آن را دستی [۵۹و۶۰] یا اتوماتیک [۶۱و۶۲] تعیین کرد. روشهای دنبال کردن با استفاده از چند دوربین مانند [۵۷و۵۸] در مواردی که با مشکل مخفی شدن شیء مواجه هستیم ، نتایج بهتری را در مقایسه با دنبال کننده‌های یک دوربینی ارائه می‌دهند. در روش مربوط به مرجع [۶۳] برای دنبال کردن ، از ترکیبی از دوربینهای ثابت و pan-tilt-zoom با زاویه دیدهایی که همپوشانی دارند استفاده شده است.

در بسیاری از موقعیتها داشتن دوربینهایی که همپوشانی دارند به دلیل محدودیت منابع یا ناحیه وسیعی که باید در حوزه دید قرار گیرد، میسر نمی‌باشد. در بعضی از این موارد باید فرضهایی در مورد سرعت و مسیر شیء برای بدست آوردن تناظر بین تصاویر مربوط به دوربینها، اتخاذ شوند [۶۴و۶۵و۶۶]. در روشهایی که تناظر را بین اشیاء در بین تصاویر گرفته شده توسط دوربینهایی که همپوشانی ندارند برقرار می‌کنند، فرض می‌شود که اولاً دوربینها ثابتند دوماً مسیر حرکت شیء مربوط به هر دوربین در دسترس است. برای مواردی که قیود مکانی _زمانی قابل استفاده نیستند برای مثال شیئی که در ناحیه غیرهمپوشانی حرکت دلخواه نامعینی دارد تنها راهکار دنبال کردن با استفاده از بازشناسی قابل اعمال است . در این راهکار هنگامی که شیء درون زاویه دید یک دوربین قرار می گیرد، از ظاهر و شکل شیء برای بازشناسی آن استفاده می‌شود.

فصل سوم

تئوری

۳-۱ مقدمه

دو روش مهم که امروزه برای دنبال کردن شیء در یک رشته ویدئویی مطرح هستند یکی روش انتقال میانگین^{۷۴} [۱ و ۲] و دیگری روش مبتنی بر فیلتر ذره ای^{۷۵} [۴ و ۵] می باشند.

در این فصل ابتدا به معرفی روش انتقال میانگین به عنوان روشی جهت یافتن مدها در فضای ویژگی می پردازیم و در قسمت بعدی، بحث استفاده از انتقال میانگین در دنبال کردن شیء را مطرح می کنیم. بعد از آن به روشهایی می پردازیم که از راهکار فضای حالت برای مدل کردن سیستم استفاده می کنند در این قسمت ابتدا فیلتر کالمن^{۷۶} [۳] 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 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. پس از مطرح کردن فیلتر ذره ای ، در بخش آخر این فصل به بررسی روشی می پردازیم که در آن مرحله ای شامل اعمال انتقال میانگین بر روی ذره ها به مراحل کار فیلتر ذره ای اضافه شده است و به دنبال کننده جدیدی به نام TSHT دست یافته است [۶].

⁷⁴ Mean Shift

⁷⁵ Particle Filter

⁷⁶ Kalman Filter

۳-۲ شیوه انتقال میانگین

این شیوه یک روش غیر پارامتریک برای آنالیز فضای ویژگی پیچیده چند مده می باشد که قادر است محدوده کلاسترهای با اشکال دلخواه را ترسیم نماید [۱]. واحد محاسباتی پایه در این روش یک شیوه تشخیص الگوی قدیمی یعنی انتقال میانگین است. به کمک این شیوه ماکزیممهای محلی (مدهای) تابع چگالی احتمال مورد جستجو قرار می گیرند.

برای معرفی شیوه انتقال میانگین ابتدا تکنیک پنجره پارزن^{۷۷} را که یکی از روشهای محبوب برای تخمین تابع چگالی می باشد را مطرح می کنیم:

در نظر بگیرید n داده $\mathbf{x}_i, i=1..n$ در فضای d بعدی داده شده است. تخمین چگالی با روش پنجره پارزن به صورت زیر به دست می آید.

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \quad (۱-۳)$$

در این رابطه K کرنل است که شکل پنجره را مشخص می کند و پارامتر h مقیاس این پنجره را تعیین می کند. کرنل K را می توان به صورت زیر تعریف کرد:

$$K(\mathbf{x}) = c_{k,d} k(\|\mathbf{x}\|^2) \quad (۲-۳)$$

⁷⁷ Parzen window

که در این رابطه $k(\mathbf{x})$ پروفایل کرنل نامیده می شود و $c_{k,d}$ ثابت نرمالیزه سازی است. رابطه ۳-۱ با توجه به

رابطه ۳-۲ به شکل زیر قابل بازنویسی می باشد:

$$\hat{f}_{h,K}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (3-3)$$

اولین گام در تحلیل فضای ویژگی با چگالی $f(\mathbf{x})$ یافتن مدهای تابع چگالی است. این مدها در محل

صفرهای گرادیان تابع چگالی قرار دارند و شیوه انتقال میانگین ، روشی برای یافتن موقعیت این صفرها بدون

تخمین تابع چگالی می باشد.

تخمین گرادیان چگالی

با فرض اینکه گرادیان تخمین تابع چگالی می تواند تخمینی برای گرادیان تابع چگالی باشد می توان رابطه

زیر را نوشت:

$$\hat{\nabla} f_{h,K}(\mathbf{x}) \equiv \nabla \hat{f}_{h,K}(\mathbf{x}) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x} - \mathbf{x}_i) k'\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (4-3)$$

با تعریف

$$g(\mathbf{x}) = -k'(\mathbf{x}) \quad (5-3)$$

$$G(\mathbf{x}) = c_{g,d} g\left(\|\mathbf{x}\|^2\right) \quad (۶-۳)$$

تخمین گرادیان تابع چگالی را می توان به شکل زیر نوشت

$$\begin{aligned} \hat{\nabla} f_{h,K}(\mathbf{x}) &= \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}) g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \\ &= \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \right] \left[\frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x} \right] \end{aligned} \quad (۷-۳)$$

در سمت راست معادله فوق بلوک اول متناسب است با تخمین چگالی در نقطه $\mathbf{x}$ که با به کارگیری کرنل G

بدست آمده است:

$$\hat{f}_{h,G}(\mathbf{x}) = \frac{c_{g,d}}{nh^d} \sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (۸-۳)$$

بلوک دوم در رابطه ۷-۳ برداری است $\mathbf{d}$ بعدی که بردار انتقال میانگین نامیده می شود.

$$m_{h,G}(\mathbf{x}) = \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x} \quad (۹-۳)$$

با توجه به تعاریف بالا گرادیان تابع چگالی به صورت زیر تخمین زده می شود:

$$\hat{\nabla} f_{h,K}(\mathbf{x}) = \hat{f}_{h,G}(\mathbf{x}) \frac{2c_{k,d}}{h^2 c_{g,d}} \mathbf{m}_{h,G}(\mathbf{x}) \quad (10-3)$$

که نتیجه می دهد

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{1}{2} h^2 c \frac{\hat{\nabla} f_{h,K}(\mathbf{x})}{\hat{f}_{h,G}(\mathbf{x})} \quad (11-3)$$

رابطه (۱۱-۳) نشان می دهد که در موقعیت $\mathbf{x}$ ، بردار انتقال میانگین که با کرنل G محاسبه شده، متناسب است با تخمین گرادیان چگالی که با کرنل K به دست آمده و توسط تخمین چگالی به دست آمده با کرنل G نرمالیزه شده است. بنابراین بردار انتقال میانگین همیشه، جهتی به سمت ماکزیمم افزایش در چگالی را دارد. از آنجا که بردار انتقال میانگین در هر نقطه از فضا هم جهت با تخمین گرادیان محلی می باشد، می تواند مسیر حرکت به سمت نقاط ایستایی تابع چگالی - که مدهای تابع چگالی جزو این نقاط هستند - را تعیین کند.

شیوه انتقال میانگین برای یافتن مد تابع چگالی با انجام پشت سر هم دو مرحله زیر انجام گیرد:

۱- بردار انتقال میانگین $\mathbf{m}_{h,G}(\mathbf{x})$ محاسبه می شود.

۲- موقعیت کرنل G به اندازه $\mathbf{m}_{h,G}(\mathbf{x})$ جابجا می شود.

این الگوریتم تا جایی ادامه می یابد که تغییرات موقعیت $\mathbf{x}$ از یک عدد کوچک که به عنوان آستانه تعیین

شده کمتر شود. این الگوریتم دارای همگرایی تضمین شده به نقطه ای است که تخمین بدست آمده از

رابطه (۳-۳) در آن گرادیان صفر دارد.

طول گام در یک الگوریتم مبتنی بر گرادیان برای کارایی ، امری بسیار مهم است . اگر طول گام خیلی بزرگ باشد، الگوریتم واگرا خواهد شد در صورتی که اگر طول گامها خیلی کوچک باشد سرعت همگرایی بسیار کند خواهد بود. روشهایی برای تعیین طول گام مناسب در الگوریتمهای گرادیانی پیشنهاد شده اند که متأسفانه هزینه بر هستند. در روش انتقال میانگین به خاطر دامنه وفقی بردار انتقال میانگین نیاز به شیوه های اضافی برای تعیین طول گام مناسب برطرف می شود این دامنه وفقی به دلیل نرمالیزه شدن بردار انتقال میانگین توسط مقدار تابع چگالی در هر نقطه ، بدست می آید و باعث می شود در نواحی با مقادیر چگالی پایین که برای تحلیل فضای ویژگی ، اهمیت زیادی ندارند طول گامها بلند باشد و در نزدیکی ماکزیممها محلی ، طول گامها کوتاه باشد و تحلیل ظریفتر صورت گیرد. بنابراین شیوه انتقال میانگین ، یک شیوه صعود گرادیان وفقی است.

برای داده های گسسته تعداد گامها برای همگرایی بستگی به کرنل استفاده شده دارند. وقتی G کرنل با تابع یکنواخت باشد همگرایی با تعداد محدودی گام بدست می آید چون تعداد موقعیت هایی که مقدار میانگین متمایزی تولید می کنند محدود است. هر چند باید این مسئله در نظر گرفته شود که حرکت در جهت گرادیان محلی در حالت کلی فقط برای گامهای بینهایت کوچک ، همگرایی را تضمین می کند.

وقتی که کرنل G وزنی را (برحسب فاصل از مرکز) به داده ها تحمیل می کند ، شیوه انتقال میانگین با بینهایت گام همگرا خواهد شد.

در این رابطه می توان به دو کرنل پر کاربرد اشاره کرد:

کرنل اپانچنیکف^{۷۸}:

که دارای پروفایل زیر است

$$k_E(x) = \begin{cases} 1-x & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases} \quad (۱۲-۳)$$

و با توجه به رابطه ۲-۳ کرنل شعاعی متقارن زیر را نتیجه می دهد

$$K_E(\mathbf{x}) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2) (1 - \|\mathbf{x}\|^2) & \|\mathbf{x}\| \leq 1 \\ 0 & otherwise \end{cases} \quad (۱۳-۳)$$

کرنل نرمال:

به شکل:

$$K_N(\mathbf{x}) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right) \quad (۱۴-۳)$$

و با پروفایل:

⁷⁸ Epanechnikov

$$k_N(x) = \exp\left(-\frac{1}{2}x\right) \quad x \geq 0 \quad (۱۵-۳)$$

تحلیل فضای ویژگی

چند مده بودن و کلاسترهای با اشکال دلخواه از مشخصه های فضاهای ویژگی واقعی می باشند. توانایی شیوه انتقال میانگین برای حرکت به سمت مد توده ای که روال از نقطه ای از آن توده آغاز شده ، آن را به یک واحد محاسباتی مناسب برای تحلیل چنین فضاهایی تبدیل کرده است. برای یافتن همه مدهای مهم ، انتقال میانگین باید بارها با نقاط آغازین مختلف اعمال شود .

۳-۳ دنبال کردن شیء با استفاده از روش انتقال میانگین

در روش دنبال کردن شیء ، مبتنی بر انتقال میانگین که در مرجع [۲] پیشنهاد شده است ، بیان شیء با استفاده از یک بیضی که محدوده شیء را تعیین می کند انجام می گیرد . ویژگی استفاد شده رنگ می باشد و در تابع مشابهت به کار رفته از متریکی که از ضریب بهاتاچاریا^{۷۹} بر گرفته شده استفاده گردیده است. و برای یافتن ناحیه ای در مجاورت محل اولیه شیء که بیشترین مشابهت را با شیء داشته باشد از روش انتقال میانگین استفاده شده است.

بیان شیء:

⁷⁹ Bhattacharyya coefficient

برای بیان شیء (داخل بیضی) از تابع چگالی احتمال در فضای ویژگی (رنگ) استفاده می شود. برای این کار ابتدا در فریم اول شیء مورد نظر توسط یک بیضی که قطر های آن موازی سطرها و ستونهای تصویر هستند به صورت دستی مشخص می شود. سپس ابعاد تصویر به گونه ای نرمالیزه می شود که بیضی مشخص کننده شیء به دایره تبدیل شود. و بعد از آن شیء مورد نظر با هیستوگرام رنگ پیکسلهای واقع شده داخل این دایره مدل می شود .

مدل تصویر هدف

پیکسلهای حاشیه دایره (در تصویر نرمالیزه شده) کمتر قابل اعتمادند زیرا بیشتر در معرض مخفی شدن پشت سایر اجسام قرار دارند و در ضمن ممکن است تحت تاثیر پس زمینه قرار گیرند بنابراین برای افزایش قدرت روش تخمین چگالی می توان از هیستوگرام وزن دار شده استفاده کرد.

وزن دار کردن هیستو گرام به این شکل صورت می گیرد که به پیکسلهای درون ناحیه دایره بیان کننده شیء، وزنهاى متفاوتی داده می شود به این صورت که به پیکسلهای مرکز دایره که از اهمیت بیشتری برخوردار هستند وزن بیشتری تخصیص داده می شود و هرچه پیکسلها از مرکز دور شوند وزن کمتری را خواهند داشت.

با فرض اینکه $\mathbf{x}_i^*$ ها، $i = 1 \dots n$ موقعیت پیکسلهای درون دایره مربوط به مدل هدف باشند و در نظر گرفتن

مرکز این دایره در مبدا مختصات، مقدار هیستوگرام در خانه u^* ام از رابطه زیر بدست می آید.

$$\hat{q}_u = C \sum_{i=1}^n k\left(\left\|\mathbf{x}_i^*\right\|^2\right) \delta[b(\mathbf{x}_i^*) - u] \quad (16-3)$$

که در آن $k(x)$ پروفایل کرنل می باشد، $b(\mathbf{x}_i^*)$ خانه پیکسلی که در موقعیت $\mathbf{x}_i^*$ قرار دارد را تعیین می

کند، m تعداد خانه های هیستوگرام می باشد و δ تابع دلتای دیراک گسسته است همچنین C ثابت

نرمالیزه سازی است که باعث برقراری شرط $\sum_{u=1}^m \hat{q}_u = 1$ می شود:

$$C = \frac{1}{\sum_{i=1}^n k\left(\left\|\mathbf{x}_i^*\right\|^2\right)} \quad (17-3)$$

بنابر این تصویر هدف با رشته $\hat{\mathbf{q}}$ که به شکل زیر تعریف شده مدل می شود.

$$\hat{\mathbf{q}} = \{\hat{q}_u\}_{u=1 \dots m} \quad (18-3)$$

کاندیدای هدف

فرض کنید در فریم جاری، $\{\mathbf{x}_i\}_{i=1 \dots n_h}$ موقعیت پیکسلهای داخل دایره ای به مرکز $\mathbf{y}$ هستند که کاندید

انطباق با مدل هدف می باشند. ناحیه کاندید را با استفاده از کرنلی مشابه کرنل استفاده شده برای تعیین

مدل شیء با عرض باند h مدل می کنیم به گونه ای که محتوای خانه u ام مربوط به ناحیه کاندید از رابطه

زیر بدست می آید:

$$\hat{p}_u(\mathbf{y}) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) \delta[b(\mathbf{x}_i) - u] \quad (۱۹-۳)$$

که در آن

$$C_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right)} \quad (۲۰-۳)$$

بنابراین ناحیه کاندید توسط رشته زیر مدل می شود:

$$\hat{\mathbf{p}}(\mathbf{y}) = \{\hat{p}_u(\mathbf{y})\}_{u=1, \dots, m} \quad \sum_{u=1}^m \hat{p}_u = 1 \quad (۲۱-۳)$$

تابع مشابهت

در گام بعدی برای مقایسه رشته های $\hat{\mathbf{p}}, \hat{\mathbf{q}}$ نیاز به یک تابع مشابهت داریم . در فریم جاری ، ماکزیمم محلی

این تابع نسبت به مختصات $\mathbf{y}$ می تواند بیانگر حضور شیء باشد. تابع مشابهت ، مشابهت بین مدل و

کاندیدای هدف را تعریف می کند. یک تابع برای تعیین فاصله بین توزیعات گسسته به شکل زیر تعریف می

گردد:

$$\hat{\rho}(\mathbf{y}) \equiv \rho[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}] = \sum_{u=1}^m \sqrt{\hat{p}_u(\mathbf{y}) \hat{q}_u} \quad (۲۲-۳)$$

این تابع بر گرفته شده از ضریب بهاتاچاریا بین $\mathbf{p}$ و $\mathbf{q}$ می باشد [۲].

مینیمم سازی فاصله :

حال باید ناحیه ای را در فریم دوم بیابیم که بیشترین مشابهت را برحسب معیار انتخاب شده با مدل شیء داشته باشد. یک راه حل جستجوی پیکسل به پیکسل کل فریم جاری می باشد، به این شکل که ناحیه بیضی کاندید را برروی کل تصویر فریم جاری بلغزانیم و هر بار مقدار تابع مشابهت را بیابیم ، مکانی که در آن تابع مشابهت ماکزیمم گردد موقعیت شیء در فریم جاری را تعیین می کند. مشکل این روش هزینه زمانی بالای آن می باشد. این مشکل را می توان با اعمال تدابیری کاهش داد مثلا به جای جستجوی کل تصویر با این فرض که شیء در دوفریم متوالی جابجایی زیادی نخواهد داشت ، جستجو را به همسایگی موقعیت شیء در فریم قبلی محدود کنیم و یا به جای حرکت پیکسل به پیکسل ، حرکت چند پیکسل به چند پیکسل را صورت دهیم. گرچه با اعمال این روشها باز هم فرایند جستجو زمانبر خواهد بود.

یکی از الگوریتمهای کارآمد برای جستجوی مشابهت ماکزیمم روش انتقال میانگین می باشد. در این روش موقعیت تصویر کاندید ، یعنی مرکز بیضی ای که مشابهت آن با تصویر مدل شیء سنجیده می شود، به صورت مکرر جابجا می شود تا جایی که در دوتکرار پشت سرهم میزان جابجایی مرکز بیضی از یک مقدار آستانه ϵ کوچکتر شود. و این نقطه همگرایی به عنوان موقعیت شیء در فریم جاری لحاظ می شود.

حداکثر کردن معیار مشابهت با استفاده از انتقال میانگین

برای رسیدن به ناحیه ای که بیشترین مشابهت با مدل را دارد، سعی می کنیم تابع $\hat{\rho}(\mathbf{y})$ را ماکزیمم کنیم.

با استفاده از دو جمله اول بسط تیلور حول مقادیر $\hat{p}_u(\hat{\mathbf{y}}_0)$ می توان رابطه ۳-۲۲ را به شکل تقریبی زیر

نوشت.

$$\rho[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{\mathbf{y}}_0) \hat{q}_u} + \frac{1}{2} \sum_{u=1}^m \hat{p}_u(\mathbf{y}) \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{\mathbf{y}}_0)}} \quad (23-3)$$

و با استفاده از رابطه ۳-۱۹ می توانیم رابطه ۳-۲۳ را به شکل زیر بازنویسی کنیم.

$$\rho[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{\mathbf{y}}_0) \hat{q}_u} + \frac{C_h}{2} \sum_{i=1}^{n_h} w_i k \left(\left\| \frac{\mathbf{y} - \mathbf{x}_i}{h} \right\|^2 \right) \quad (24-3)$$

که در آن

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{\mathbf{y}}_0)}} \delta[b(\mathbf{x}_i) - u] \quad (25-3)$$

ترم دوم سمت راست رابطه (۳-۲۴) بیانگر تخمین چگالی محاسبه شده به کمک پروفایل کرنل k در محل $\mathbf{y}$

می باشد، که با ضریب w_i متناظر با هر داده، وزن دار شده است. با توجه به ثابت بودن قسمت اول سمت

راست معادله ۳-۲۴، مد این چگالی مذکور در یک همسایگی محلی، همان ماکزیمم مورد نظر ماست. این

مد با اعمال شیوه انتقال میانگین [۱] قابل تحصیل است. برای به دست آوردن این مد ، کرنل به صورت

تکراری از موقعیت قبلی $\hat{\mathbf{y}}_0$ به موقعیت جدیدش $\hat{\mathbf{y}}_1$ طبق رابطه زیر جابجا می شود:

$$\hat{\mathbf{y}}_1 = \frac{\sum_{i=1}^{n_h} \mathbf{x}_i w_i g\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right)} \quad (۲۶-۳)$$

که در آن $g(\mathbf{x}) = -k'(\mathbf{x})$. این کار تا جایی که $\|\hat{\mathbf{y}}_1 - \hat{\mathbf{y}}_0\| < \varepsilon$ شود، ادامه می یابد. می توان علاوه بر این

قید ، حد بالایی برای تعداد تکرارها نیز تعیین کرد.

برای بهبود عملکرد و همچنین گسترش حوزه کاربرد روش فوق می توان از روشهای تخمین حرکت و

تخصیص داده در کنار این روش استفاده کرد.

حوزه عملکردی جذب^{۸۱}

یک مسئله مهم در روش دنبال کردن مبتنی بر انتقال میانگین این است که برای عملکرد صحیح الگوریتم ،

نقطه ای که انتقال میانگین از آن نقطه کار خود را شروع می کند تا مرکز واقعی شیء چه مقدار می تواند

فاصله داشته باشد. ناحیه ای در فریم جاری که موقعیت جدید شیء در آن توسط الگوریتم دنبال کننده

مبتنی بر انتقال میانگین قابل پیدا کردن است را حوزه عملکردی جذب می نامیم. در واقع برای هر مد تابع

⁸¹ Operational basin of attraction

چگالی احتمال یک حوزه عملکردی جذب وجود دارد که این حوزه ناحیه ای است که اگر از هر نقطه از آن فرایند انتقال میانگین را آغاز کنیم ، الگوریتم به آن مد همگرا خواهد شد.

در دنبال کننده انتقال میانگین برای هر نقطه همگرایی (مد تابع چگالی احتمال) ، به دلیل استفاده از کرنل، حوزه جذب الزاما شامل ناحیه ای به اندازه شیء و به مرکزیت آن نقطه همگرایی می باشد. این ناحیه را حوزه قطعی جذب می نامیم. بنابراین می توان گفت اگر نقطه آغاز انتقال میانگین که همان نقطه همگرایی انتقال میانگین در فریم قبل می باشد در فریم جاری در ناحیه ای قرار گیرد که توسط شیء مورد پوشش قرار گرفته است ، نقطه همگرایی انتقال میانگین نشانگر قابل اعتمادی برای موقعیت جدید شیء می باشد. در اینجا فرض بر این است که بیان شیء، تمایز کافی را برای انطباق کاندید و مدل ایجاد می کند به گونه ای که در همسایگی موقعیت مرکز شیء ، یک ماکزیمم منحصر به فرد برای معیار مشابهت وجود داشته باشد.

۳-۴ 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:

$$\mathbf{x}_k = \begin{bmatrix} x_k \\ y_k \\ v_{xk} \\ v_{yk} \end{bmatrix} \quad (3-27)$$

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 $\mathbf{z}_k$ مشخص می شوند:

$$\mathbf{z}_k = \begin{bmatrix} xm_k \\ ym_k \end{bmatrix} \quad (28-3)$$

عناصر این بردار موقعیت شیء در فریم جاری می باشند که توسط یک روش آشکارسازی 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 ^{۸۲} پروسه و معادله سنجش) را در اختیار داریم:

$$\mathbf{x}_{k+1} = F_{k+1,k} \mathbf{x}_k + \mathbf{w}_k \quad (29-3)$$

که معادله پروسه نامیده می شود. در این معادله $F_{k+1,k}$ ماتریس انتقال ^{۸۳} و $\mathbf{w}_k$ 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 ^{۸۴} 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:

⁸² Process equation

⁸³ Transition matrix

⁸⁴ Measurement equation

$$\mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{v}_k \quad (30-3)$$

در این رابطه $\mathbf{z}_k$ مشاهده ما در زمان k و H_k ماتریس سنجش^{۸۵} و $\mathbf{v}_k$ 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 $p(\mathbf{x}_0 | \mathbf{z}_0)$ به صورت گوسی باشد همچنین نویزهای سیستم و اندازه گیری با توزیع گوسی فرض شوند و توابع پیاده سازی شده با ماتریسهای $F_{k+1,k}$ و H_k توابعی خطی باشند.

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

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ v_{xk+1} \\ v_{yk+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ v_{xk} \\ v_{yk} \end{bmatrix} + \mathbf{w}_k \quad (31-3)$$

$$\begin{bmatrix} xm_k \\ ym_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ v_{xk} \\ v_{yk} \end{bmatrix} + \mathbf{v}_k \quad (32-3)$$

⁸⁵ Measurement matrix

روش کار فیلتر کالمن را می توان در دو مرحله بیان کرد [۳] 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 $\hat{\mathbf{x}}_k^-$

نمایش می دهیم و از رابطه زیر بدست می آید:

$$\hat{\mathbf{x}}_k^- = F_{k,k-1} \hat{\mathbf{x}}_{k-1} \quad (33-3)$$

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

دو گام فوق را می توان با رابطه زیر نمایش داد

$$\hat{\mathbf{x}}_k = G_k^{(1)} \hat{\mathbf{x}}_k^- + G_k \mathbf{z}_k \quad (34-3)$$

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

با استفاده از روابط ریاضی می توان نشان داد [۳]:

$$G_k^{(1)} = I - G_k H_k \quad (35-3)$$

بنابراین رابطه ۳۴-۳ به شکل زیر قابل بیان می باشد:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + G_k (\mathbf{z}_k - H_k \hat{\mathbf{x}}_k^-) \quad (36-3)$$

که در آن ماتریس G_k بهره کالمن^{۸۶} نامیده می شود. می توان نشان داد [۳]:

⁸⁶ Kalman gain

$$G_k = P_k^- H_k^T [H_k P_k^- H_k^T + R_k]^{-1} \quad (37-3)$$

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

$$P_k^- = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)^T] \quad (38-3)$$

رابطه ۳۷-۳ بهره کالمن را که برای حل معادله (۳۶-۳) با فرض داشتن حالت قبلی، تنها فاکتور مجهول می

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

باشد که مجهول است. با توجه به اینکه $\mathbf{x}_k$ رادر دسترس نداریم نمی توانیم P_k^- را مستقیماً محاسبه کنیم

لذا برای تکمیل عملکرد فیلتر از یک روش باز گشتی استفاده می کنیم. این روش شامل دو مرحله می باشد:

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

می کنیم:

$$P_k^- = F_{k,k-1} P_{k-1} F_{k,k-1}^T + Q_{k-1} \quad (39-3)$$

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

$$P_k = (I - G_k H_k) P_k^- \quad (40-3)$$

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

$$\hat{\mathbf{x}}_0 = E[\mathbf{x}_0]$$
$$P_0 = E[(\mathbf{x}_0 - E[\mathbf{x}_0])(\mathbf{x}_0 - E[\mathbf{x}_0])^T]$$

←

تعیین اطلاعات اولیه



پیش بینی حالت

$$\hat{\mathbf{x}}_k^- = F_{k,k-1} \hat{\mathbf{x}}_{k-1}$$



$$P_k^- = F_{k,k-1} P_{k-1} F_{k,k-1}^T + Q_{k-1}$$



$$G_k = \frac{P_k^- H_k^T}{H_k P_k^- H_k^T + R_k}$$



به روز کردن

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



$$P_k = (I - G_k H_k) P_k^-$$



۳-۵ فیلتر ذره ای

همانطور که در بخش ۳-۴ گفته شد برای استفاده از فیلتر کالمن باید چند شرط محدود کننده را در نظر بگیریم به این شکل که تابع چگالی احتمال اولیه $p(\mathbf{x}_0 | \mathbf{z}_0)$ باید به صورت گوسی باشد همچنین باید نویزهای سیستم و اندازه گیری با توزیع گوسی فرض شوند و توابع پیاده سازی شده با ماتریسهای $F_{k+1,k}$ و H_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 [۴] استفاده کرد. فیلتر ذره ای تکنیکی برپایه 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 کارکردن با مدل های غیر خطی و امکان پیاده سازی آسان و کارآمد.

در اینجا نیز مشابه قسمت ۳-۴ فرض میکنیم معادله پروسه و معادله سنجش را در اختیار داریم این معادلات را این بار به صورت زیر مطرح می کنیم

معادله پروسه :

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{v}_k) \quad (41-3)$$

و معادله سنجش :

$$\mathbf{z}_k = h_k(\mathbf{x}_k, \mathbf{n}_k) \quad (42-3)$$

که در این روابط h_k و f_k توابع غیرخطی و متغیر با زمان می باشند و $\mathbf{v}_k$ و $\mathbf{n}_k$ 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 $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ مطرح کرد. در این رابطه فرض می شود که تابع احتمال مربوط به حالت اولیه سیستم در اختیار باشد:

$$p(\mathbf{x}_0 | \mathbf{z}_0) = p(\mathbf{x}_0) \quad (43-3)$$

سپس دو گام پیش بینی و به روز کردن برای بدست آوردن تابع چگالی احتمال $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ به کار گرفته می شود فرض کنید $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ در زمان $k-1$ در دسترس است. در این مرحله پیش بینی با استفاده از معادله حالت به شکل زیر صورت می گیرد

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \quad (44-3)$$

با داشتن $p(\mathbf{x}_k | \mathbf{z}_{1:k-1})$ و با استفاده از قاعده بیز داریم :

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})} \quad (45-3)$$

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 $\{\mathbf{s}_k^{(n)}\}_{n=1,2,\dots,N_s}$ 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:

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) \approx \sum_{n=1}^{N_s} \pi_k^{(n)} \delta(\mathbf{x}_k - \mathbf{s}_k^{(n)}) \quad (46-3)$$

که در آن وزن $\pi_k^{(n)}$ عددی متناسب با $p(\mathbf{z}_k | \mathbf{s}_k^{(n)})$ در نظر گرفته می شود.

در اینجا الگوریتم موسوم به انتشار چگالی شرطی ^{۸۷} [۵] که یکی از مشهورترین روشهای پیاده سازی فیلتر ذره ای در زمینه بینایی ماشین می باشد را مطرح می کنیم:

تکرار

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 شکل 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 $p(\mathbf{x}_k | \mathbf{x}_{k-1} = s_k'^{(n)})$ انجام بده.

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

$$\pi_k^{(n)} = p(\mathbf{z}_k | \mathbf{x}_k = s_k'^{(n)})$$

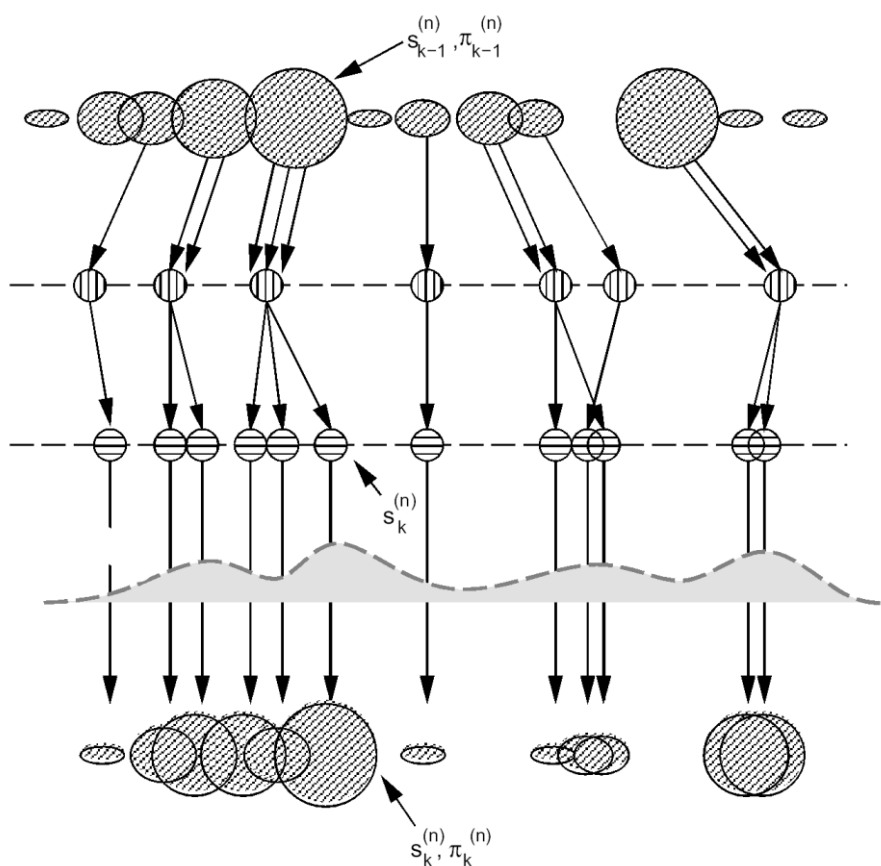
سپس به گونه ای نرمالیزه کن که $\sum_n \pi_k^{(n)} = 1$ و در کنار احتمال تجمعی به شکل $\{s_k^{(n)}, \pi_k^{(n)}, c_k^{(n)}\}$ قرار بده که در آن

$$c_k^{(0)} = 0,$$

$$c_k^{(n)} = c_k^{(n-1)} + \pi_k^{(n)} \quad (n=1, \dots, N)$$

الگوریتم ۳-۱: الگوریتم انتشار چگالی شرطی [۵]

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 $p(\mathbf{z}_{k-1} | \mathbf{s}_{k-1}^{(n)})$ 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 $(\mathbf{z}_k)$ وزن دهی می شود (سطر آخر در شکل) و در نهایت در محل هر نمونه احتمال پسین با توجه به معادله (۳-۴۶) قابل محاسبه خواهد بود.



شکل ۳-۲ عملکرد الگوریتم انتشار چگالی شرطی

فیلتر ذره ای مسئله تخمین حالت غیر گوسی و غیر خطی 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 ره می باشد. ه این تعداد با افزایش ابعاد فضای حالت به شدت افزایش می یابد و باعث افزایش هزینه محاسباتی می گردد.

۳-۶ دنبال کننده TSHT

دنبال کننده TSHT [۶] عملکردی مشابه الگوریتم فیلتر ذره ای دارد با این تفاوت که یک مرحله به آن افزوده شده است. در TSHT پس از اعمال معادله پروسه و قبل از وزن دهی نمونه ها ، یک مرحله به شکل زیر اضافه می شود.

در این مرحله بر روی تک تک ذره ها عملیات انتقال میانگین صورت می گیرد و هر یک از ذره ها به محل همگرایی مربوطه منتقل می شوند. در روش مذکور وزن هر ذره با استفاده از ضرائب بهاتاچاریا تعیین می شود. و موقعیت نهایی که به عنوان محل شیء معرفی می شود از رابطه زیر بدست می آید.

$$E[\mathbf{x}_t | \mathbf{z}_{1:t}] \approx \frac{1}{N_s} \sum_{n=1}^{N_s} \pi_k^{(n)} \mathbf{x}_k^{(n)} \quad (47-3)$$

همچنین در این روش پروسه زیر به کار گرفته شده است:

$$\mathbf{x}_k = \mathbf{x}_{k-1} + C_k \mathbf{v}_k \quad (48-3)$$

که در آن

$$C_k \propto \frac{1}{r} \sum_{n=k-r}^k |\mathbf{x}_n - \mathbf{x}_{n-1}| \quad (49-3)$$

و بردار تصادفی $\mathbf{v}_k = N(0, I)$ دارای توزیع گوسی با میانگین صفر و ماتریس کواریانس یکانی می باشد.

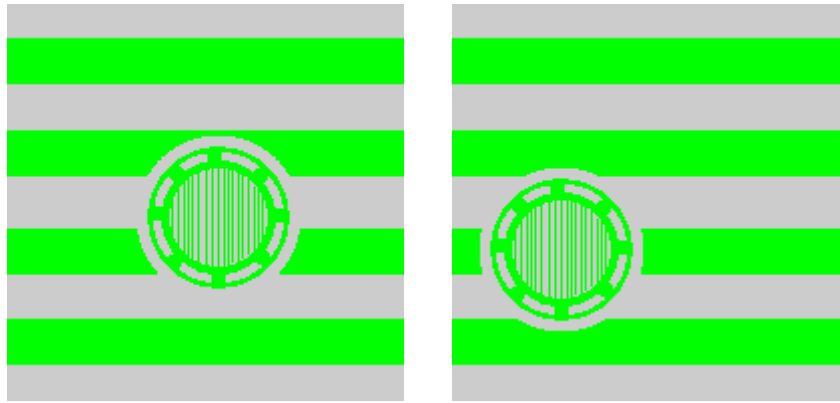
فصل چهارم

روشهای پیشنهادی

در این فصل جهت بهبود روشهای دنبال کردن مبتنی بر انتقال میانگین و فیلتر ذره ای دو روش را پیشنهاد می کنیم. اولین روش مربوطه به استفاده از بیان جدیدی برای شیء در دنبال کننده انتقال میانگین است که در این روش برای بیان شیء از اطلاعات جهت و دامنه گرادیان کمک گرفته می شود. این تغییر منجر به بهبود دنبال کنندگی شیء در مواردی که رنگهای به کار رفته در شیء و پس زمینه مشابه هستند می شود. دومین روش پیشنهاد شده در این فصل شامل اضافه کردن مرحله ای برگرفته از روش انتقال میانگین به فیلتر ذره ای می باشد. این روش قادر به دنبال کردن اشیاء سریع می باشد و سرعت اجرای آن نیز خیلی بیشتر از سرعت اجرای فیلتر ذره ای است.

۴-۱ استفاده از اطلاعات جهت و دامنه گرادیان در دنبال کننده انتقال میانگین

یک دنبال کننده مبتنی بر انتقال میانگین که از هیستوگرام رنگ استفاده می کند [۲]، در مواردی که رنگهای موجود در شیء و پس زمینه مشابه می باشند، ممکن است شیء را گم کند. هنگامی که تنها از هیستوگرام رنگ برای بیان شیء استفاده می کنیم اطلاعات مکانی مربوط به حضور رنگ های موجود در شیء را از دست می دهیم. برای مثال در شکل (۴-۱الف) شیء متحرک که دایره با حاشیه خاکستری است، بافتی کاملاً متفاوت با پس زمینه دارد ولی رنگهای آن دو مشابهند. در صورت جابجایی شیء (شکل ۴-۱ب)، دنبال کننده ای که تنها از هیستوگرام رنگ استفاده می کند با مشکل مواجه خواهد شد.

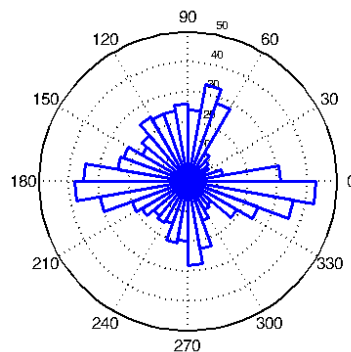


ب

الف

شکل ۴-۱ الف) شیء در پس زمینه ای با رنگهای مشابه ب) جابجایی شیء

برای رفع این مشکل پیشنهاد می کنیم تا از اطلاعات جهت و دامنه گرادیان پیکسلها علاوه بر رنگ استفاده شود. هیستوگرام جهت گرادیان به تنهایی در کاربردهای مختلفی در بینایی ماشین مورد استفاده قرار گرفته است مانند [۶۸و۶۷]. به عنوان مثالی برای هیستوگرام جهت گرادیان، شکل (۴-۲ب) را در نظر بگیرید، در این شکل هیستوگرام جهت گرادیان مربوط به تصویر شخص در حال قدم زدن در شکل (۴-۲الف)، نمایش داده است [۶۸]. همانطور که در این دیاگرام نشان داده شده است فضای هیستوگرام با ۳۶ خانه، افراز شده است.



ب



الف

شکل ۴-۲ نمونه ای از هیستوگرام جهت گرادیان (IEEE ۲۰۰۵ ©) الف) تصویر شخص در حال قدم زدن ب) نمایش هیستوگرام جهت گرادیان به صورت قطبی

در روش پیشنهادی ما ، علاوه بر اطلاعات جهت گرادیان اطلاعات دامنه گرادیان پیکسلهای تصویر نیز به عنوان یک ویژگی مستقل به کارگرفته می شود. ضمن اینکه از اطلاعات رنگ که می تواند بسیار ارزشمند باشد نیز صرف نظر نمی کنیم . برای استفاده همزمان از این اطلاعات، فضای هیستوگرام را با پنج بعد مستقل (شامل b, g, r ، جهت گرادیان و دامنه گرادیان) شکل می دهیم و این فضا را به خانه هایی که آن را افراز می کنند، تقسیم می کنیم و سپس هر پیکسل از تصویر را با توجه به مولفه های رنگ ، دامنه گرادیان و جهت گرادیان آن پیکسل به خانه مربوطه متناظر میکنیم . در این روش برای بدست آوردن دامنه گرادیان هر پیکسل تصویر از رابطه زیر استفاده می کنیم.

$$m(x, y) = \sqrt{(I(x+1, y) - I(x-1, y))^2 + I(x, y+1) - I(x, y-1))^2} \quad (۱-۴)$$

که در آن $I(x, y)$ شدت روشنایی تصویر در موقعیت (x, y) را مشخص می کند .

و برای بدست آوردن جهت گرادیان هر پیکسل تصویر از رابطه زیر استفاده می کنیم.

$$\theta(x, y) = \tan^{-1}((I(x, y+1) - I(x, y-1)) / (I(x+1, y) - I(x-1, y))) \quad (۲-۴)$$

که با تعریف

$$a = I(x, y+1) - I(x, y-1) \quad (۳-۴)$$

9

$$b = I(x+1, y) - I(x-1, y) \quad (۴-۴)$$

می توان رابطه (۲-۴) را به صورت زیر باز نویسی کرد

$$\theta(x, y) = \tan^{-1}(a/b) \quad (۵-۴)$$

در رابطه فوق θ عددی بین -90 تا 90 درجه بدست خواهد آمد ضمن اینکه با این مشکل مواجه هستیم که

در صورتی که $a=b=0$ شود θ قابل محاسبه نخواهد بود برای تعیین زاویه در محدوده $[0, 360]$ و حل این

مشکل به شیوه زیر عمل می کنیم:

ابتدا b را مورد بررسی قرار می دهیم:

$$\text{if } b = 0 \rightarrow \begin{cases} a < 0 \rightarrow \theta = 180 \\ a > 0 \rightarrow \theta = 0 \\ a = 0 \rightarrow \theta = ND \text{ (not Defined)} \end{cases}$$

و اگر b مخالف صفر بود آنگاه مقدار θ را بررسی میکنیم

$$\text{if } \theta \in (-90, 0) \rightarrow \begin{cases} a > 0 \rightarrow \theta = \theta + 90 \\ a < 0 \rightarrow \theta = \theta + 270 \end{cases}$$

$$\text{if } \theta \in (0,90) \rightarrow \begin{cases} a > 0 \rightarrow \theta = \theta + 270 \\ a < 0 \rightarrow \theta = \theta + 90 \end{cases}$$

$$\text{if } \theta = 0 \rightarrow \begin{cases} b < 0 \rightarrow \theta = 90 \\ b > 0 \rightarrow \theta = 270 \end{cases}$$

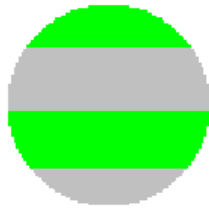
در روش پیشنهادی برای تعیین هیستوگرام جهت گرادیان علاوه بر تقسیم بندی محدوده $[0, 360]$ ، یک قسمت هم برای پیکسلهایی که در آنها $\theta = ND$ است (زاویه نامفهوم می باشد)، در نظر می گیریم زیرا هنگام محاسبه وزن مربوط به هر پیکسل با استفاده از رابطه (۳-۲۵) لازم است هر پیکسل، متناظر با یک خانه در فضای هیستوگرام باشد.

در این روش برای بدست آوردن مدل تصویر هدف و تصویر کاندید (مانند [۲]) از روابط ۳-۱۶ تا ۳-۲۱ استفاده می کنیم و تنها تفاوت در تابع $b(x_i)$ می باشد به این شکل که در اینجا $b(x_i)$ تابعی از g, b, r ، جهت گرادیان و دامنه گرادیان در موقعیت x_i می باشد:

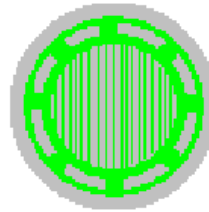
$$b(x_i) = f(\theta(x_i), m(x_i), R(x_i), G(x_i), B(x_i)) \quad (۴-۶)$$

در حالی که در مرجع [۲]، $b(x_i)$ تنها تابعی از g, b, r در نظر گرفته شده است.

در شکل ۴-۳ تصویر شیء همچنین قسمتی از پس زمینه مربوط به شکل ۴-۱ آورده شده است.



ب



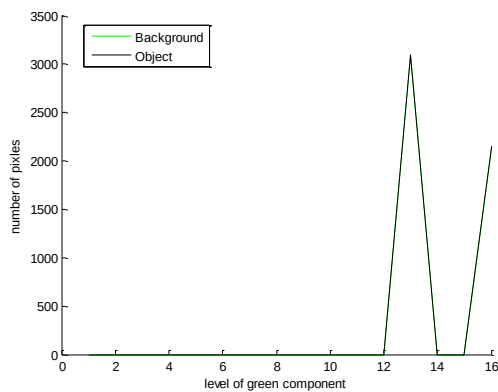
الف

شکل ۳-۴ (الف) تصویر شیء در شکل ۱-۴ (ب) تصویر بخشی از پس زمینه مربوط به شکل ۱-۴

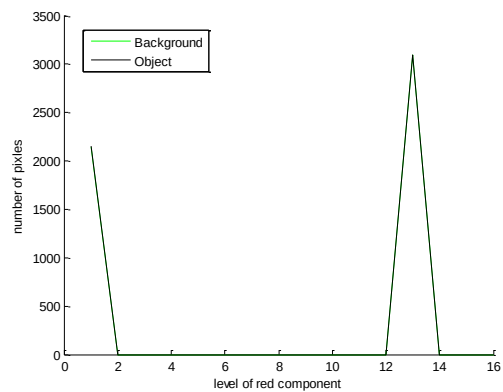
در شکل ۴-۴ برای مقایسه مولفه های b, g, r ، دامنه گرادیان و جهت گرادیان مربوط به تصاویر ۳-۴ الف و ۴-

۳ ب ، هیستوگرامهای مربوط به این مولفه ها به صورت مجزا از یکدیگر و به صورت هیستوگرام یک بعدی

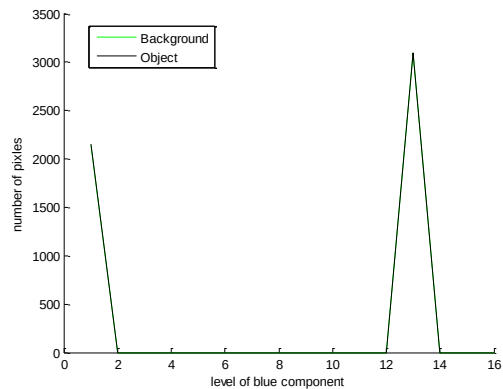
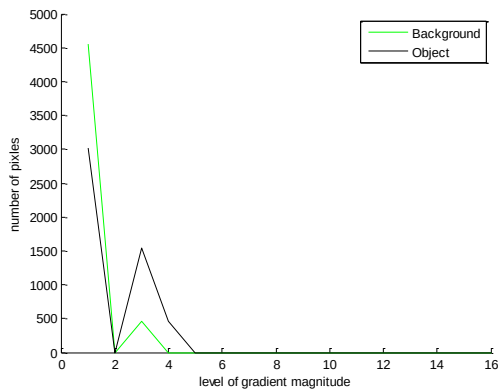
نمایش داده شده است.



ب

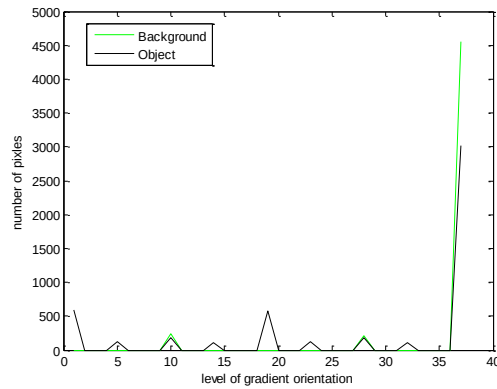


الف



د

ج



ه

شکل ۴-۴ مقایسه هیستوگرام تصاویر ۳-۴ الف(شیء) و ۳-۴ ب(پس زمینه): الف ، ب ، ج) به ترتیب مولفه های رنگ b, g, r که در هر سه مورد دو نمودار نمایش داده شده به قدری به هم شبیه هستند که به صورت یک نمودار دیده می شوند. د) دامنه گرادیان ه) جهت گرادیان

برای تشکیل این هیستوگرامها مولفه های b, g, r و دامنه گرادیان هر یک به ۱۶ قسمت و جهت گرادیان به

۳۶ قسمت تقسیم بندی شدند و برای پیکسلهایی که در آنها $\theta = ND$ می شود ، قسمت ۳۷ ام در نظر

گرفته شده است. کرنل به کار گرفته شده نیز کرنل یکنواخت می باشد.

همان طور که مشاهده می شود هیستوگرامهای b, g, r بسیار شبیه به هم هستند به گونه ای که در شکل

های مربوطه به صورت یک نمودار دیده می شوند در حالی که در مورد هیستوگرامهای دامنه گرادیان و

جهت گرادیان ، تفاوت قابل ملاحظه ای بین دو نمودار رسم شده دیده می شود.

۴-۲ استفاده از روش انتقال میانگین در فیلتر ذره‌ای

در مورد دنبال کردن به روش انتقال میانگین همانگونه که در بخش ۳-۳ اشاره کردیم در هر فریم اگر نقطه آغاز انتقال میانگین ، در ناحیه ای قرار گیرد که توسط شیء مورد پوشش قرار گرفته است ، نقطه همگرایی الگوریتم انتقال میانگین ، موقعیت شیء را نشان خواهد داد .

از این مسئله می توان نتیجه گرفت از هر نقطه ای در ناحیه ای به اندازه شیء و به مرکزیت یک نقطه همگرایی (یعنی حوزه قطعی جذب آن نقطه همگرایی)، انتقال میانگین را آغاز کنیم الگوریتم به همان نقطه همگرایی، همگرا خواهد شد. البته این با فرض این است که بیان شیء، تمایز کافی را برای انطباق کاندید و مدل ایجاد می کند به گونه ای که در همسایگی موقعیت مرکز شیء ، یک ماکزیمم منحصر به فرد برای معیار مشابهت وجود داشته باشد.

روش پیشنهادی ما مشابه روش TSHT [۶] است با این تفاوت که پس از انجام عملیات انتقال میانگین بر روی یک ذره و رسیدن به نقطه همگرایی ، ناحیه ای به اندازه شیء و به مرکزیت این نقطه همگرایی در نظر گرفته می شود و فرض می شود هر ذره ای که در این محدوده قرار دارد به همان نقطه همگرایی، همگرا خواهد شد . علاوه بر این فرض می شود اگر ذره ای در حین تکرارهای انتقال میانگین وارد این حوزه شود ، به نقطه همگرایی تعیین شده، همگرا خواهد شد و اعمال انتقال میانگین بر روی آن ذره متوقف می شود . به عبارت دیگر در این روش نتیجه اعمال انتقال میانگین بر روی ذره هایی که از ابتدا (یا در طول عملیات انتقال میانگین) در حوزه جذب قطعی یک نقطه همگرایی قرار گیرند ، بدون انجام انتقال میانگین (یا ادامه آن) ، همان نقطه همگرایی تعیین می شود که این کار منجر به کاهش زمان لازم برای اجرای الگوریتم می شود.

الگوریتم پیشنهادی در واقع همان الگوریتم انتشار چگالی شرطی [۵] است که بعد از مرحله دوم آن ، یک

مرحله برای اعمال انتقال میانگین اضافه شده است. الگوریتم ۴-۱ روش پیشنهادی را بیان می کند.

تکرار

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 شکل زیر انتخاب کن:

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

(ب) کوچکترین j را که به ازای آن $c_{k-1}^{(j)} \geq r$ 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 $p(\mathbf{x}_k | \mathbf{x}_{k-1} = s_k'^{(n)})$ انجام بده.

۳- اعمال انتقال میانگین بر روی نمونه

حلقه:

(الف) موقعیت نمونه را بررسی کن اگر در حوزه جذب نقطه همگرایی بدست آمده از اعمال

انتقال میانگین بر روی یکی از نمونه های قبلی قرار دارد ، نمونه را در موقعیت آن نقطه

همگرایی قرار بده و از حلقه خارج شو.

(ب) تکرار دیگری از انتقال میانگین را انجام بده و نمونه را در نقطه بدست آمده قرار بده .

(ج) شرط تعیین شده برای پایان انتقال میانگین را (جابجایی کوچکتر از مقدار آستانه و یا

رسیدن به حداکثر تکرار) بررسی کن اگر شرط محقق نشده بود به (الف) برو و اگر محقق شده

الگوریتم ۱-۴ روش پیشنهادی برای استفاده از روش انتقال میانگین در فیلتر ذره ای

فصل پنجم

نتایج عملی

در این فصل به ارزیابی عملکرد روشهای پیشنهادی در فصل چهارم می پردازیم. در ادامه نتیجه آزمایشات انجام شده در این زمینه و تحلیل بعضی از نتایج آورده شده است .

۵-۱ استفاده از اطلاعات جهت و دامنه گرادیان در دنبال کننده انتقال

میانگین

دنبال کننده انتقال میانگین را یک بار با استفاده از هیستوگرام رنگ (همانند مرجع [۲]) و بار دیگر با به کارگیری هیستوگرام پیشنهادی (هیستوگرام مبتنی بر رنگ، دامنه گرادیان و جهت گرادیان)، بر روی دو رشته ویدئویی که در ادامه آنها را با رشته های ۱ و ۲ نام می بریم ، اعمال کردیم.

در تصاویر ۵-۱ تا ۵-۴ نتایج این پیاده سازی ها برای چند فریم از هر رشته آورده شده است. در این تصاویر دایره سفید رنگ بیانگر ناحیه ای است که دنبال کننده آن را به عنوان ناحیه شیء تعیین کرده است (به غیر از فریم اول در هر رشته که در آن موقعیت شیء مورد نظر به صورت دستی تعیین شده است). همانطور که در تصاویر ۵-۱ و ۵-۳ نشان داده شده است هنگامی که در پس زمینه، رنگهای مشابه با رنگهای موجود در تصویر شیء وجود دارد دنبال کننده انتقال میانگین مبتنی بر هیستوگرام رنگ، در پیدا کردن شیء دچار مشکل می شود و در نهایت شیء را گم می کند در حالی دنبال کننده مبتنی بر هیستوگرام پیشنهادی قادر به دنبال کردن شیء می باشد (تصاویر ۵-۲ و ۵-۴).

در پیاده سازی روش انتقال میانگین مبتنی بر هیستوگرام رنگ، برای تشکیل هیستوگرام از فضای رنگ **RGB** با $16 \times 16 \times 16$ خانه استفاده گردید. همچنین برای تشکیل هیستوگرام پیشنهادی هر بعد از عناصر رنگ به ۴ قسمت ، بعد مربوط به دامنه گرادیان به ۱۶ قسمت و بعد مربوط به جهت گرادیان به ۳۶ قسمت

تقسیم شد که با توجه به قسمت مربوط به پیکسل‌های با $\theta = ND$ (پیکسل‌هایی که در آنها گرادیان دارای

جهت نا مفهوم می باشد)، تعداد خانه ها $(36+1) \times 16 \times 4 \times 4 \times 4$ می باشد.

در پیاده سازی روشهای مذکور مدل شیء مطابق با ناحیه انتخابی در فریم اول می باشد و بروزرسانی نمی

شود.



ج

ب

الف



و

د

شکل ۵-۱: رشته ۱، روش انتقال میانگین با استفاده از هیستوگرام رنگ

الف: فریم ۱ ، ب: فریم ۱۴ ، ج: فریم ۱۷ ، د: فریم ۲۶ ، و: فریم ۳۳



ج

ب

الف

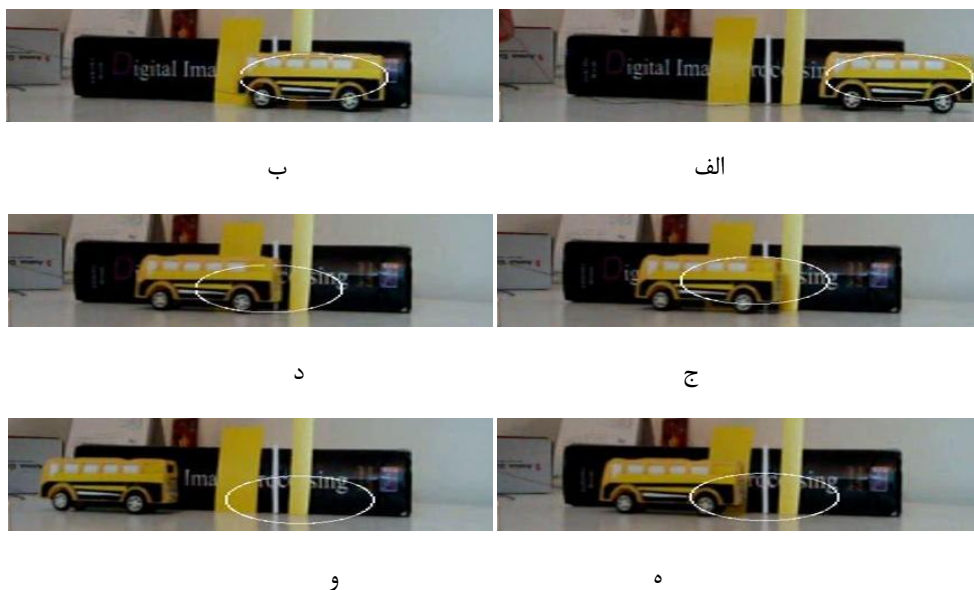


و

د

شکل ۵-۲: رشته ۱، روش انتقال میانگین با استفاده از هیستوگرام پیشنهادی

الف: فریم ۱ ، ب: فریم ۱۴ ، ج: فریم ۱۷ ، د: فریم ۲۶ ، ه: فریم ۳۳



شکل ۵-۳: رشته ۲، روش انتقال میانگین با استفاده از هیستوگرام رنگ

الف: فریم ۱ ، ب: فریم ۵۲ ، ج: فریم ۷۹ ، د: فریم ۸۲ ، ه: فریم ۹۰ ، و: فریم ۱۱۰

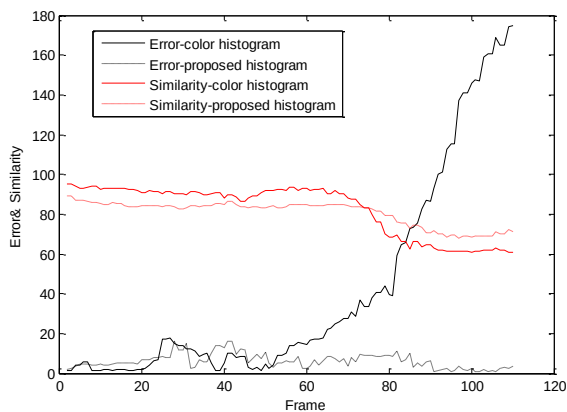


شکل ۵-۴: رشته ۲، روش انتقال میانگین با استفاده از هیستوگرام پیشنهادی

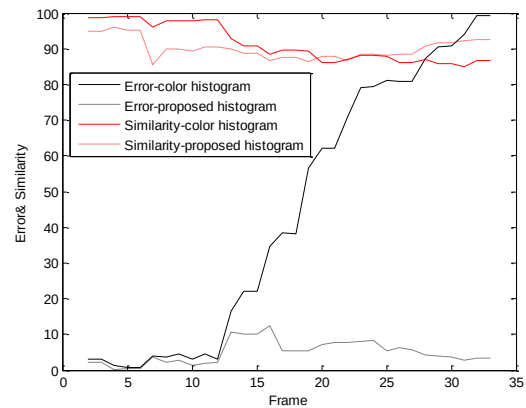
الف: فریم ۱ ، ب: فریم ۵۲ ، ج: فریم ۷۹ ، د: فریم ۸۲ ، ه: فریم ۹۰ ، و: فریم ۱۱۰

در دو رشته مورد آزمایش برای ارزیابی کمی خطای دنبال کننده ها، مکان واقعی شیء در هر فریم به صورت دستی تعیین گردید. خطا در پیدا کردن موقعیت شیء در هر فریم مطابق نمودارهای شکل (۵-۵) می باشد.

در این نمودارها خطا با استفاده از رابطه $Error = \sqrt{We^2 + He^2}$ محاسبه شده است که در آن We خطای افقی و He خطای عمودی موقعیت شیء بر حسب تعداد پیکسل هستند.



ب



الف

شکل ۵-۵: خطا در یافتن محل شیء به همراه مشابهت بین مدل شیء و مدل ناحیه ای که در هر فریم توسط الگوریتم به عنوان شیء تعیین شده است
(الف) رشته ۱ (ب) رشته ۲

هم چنین در شکل (۵-۵) مشابهت بین ناحیه ای که توسط الگوریتم دنبال کننده شیء در هر فریم به عنوان شیء تعیین شده است با مدل شیء آورده شده است. در این نمودار ها مشابهت با استفاده از رابطه (۳-۲۲) محاسبه گردیده و به در صد بیان شده است.

همانطور که در نمودار شکل (۵-۵الف) نمایش داده شده هنگامی که تنها اطلاعات رنگ را به کار می بریم ، علیرغم افزایش خطا ، میزان مشابهت هیستوگرامهای رنگ در مدل و خروجی دنبال کننده بالا می باشد همچنین درمورد رشته دو نیز با مقایسه نمودارهای مربوط به حالت استفاده از رنگ در شکل (۵-۵ب) می توان گفت در هنگامی که خطا افزایش می یابد مشابهت همچنان زیاد است و تا چند فریم افزایش خطا در حالی روی می دهد که ناحیه پیدا شده توسط الگوریتم، مشابهت زیادی با مدل شیء دارد، این نشان دهنده نا مناسب بودن شیوه بیان شیء (استفاده از هیستوگرام رنگ) می باشد.

برای ارزیابی بهتر الگوریتمهای موردنظر ، پیاده سازی های مشابهی بر روی دو رشته دیگر که آنها را رشته های ۴۰۳ می نامیم (شکل ۵-۶ الف و ب) صورت گرفت ، در این دو رشته هر دو روش قادر به دنبال کردن شیء می باشند. از نتایج این پیاده سازی ها برای مقایسه سرعت اجرای دو روش استفاده نموده ایم. ضمن اینکه با استفاده از موقعیت واقعی شیء که به طور دستی محاسبه شده است ، خطای این دو روش نیز مورد مقایسه قرار گرفتند . نتایج پیاده سازی الگوریتم دنبال کننده مبتنی بر انتقال میانگین بر روی رشته های ۴۰۳ نشان می دهد خطای دنبال کننده مبتنی بر هیستوگرام رنگ برای رشته سه ۰/۹۹ برابر و برای رشته چهار ۱/۸۱ برابر خطای دنبال کننده مبتنی بر هیستوگرام پیشنهادی می باشد. همچنین زمان لازم برای

پیاده سازی دنبال کننده مبتنی بر هیستوگرام رنگ برای رشته سه ۰/۸۹ برابر و برای رشته چهار ۰/۷۱ برابر زمان لازم برای پیاده سازی دنبال کننده مبتنی بر هیستوگرام پیشنهادی می باشد.

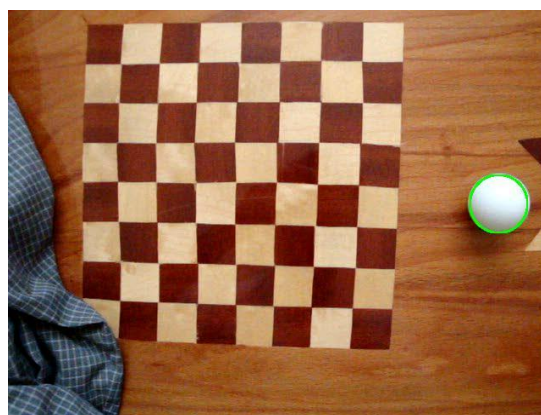
۵-۲ استفاده از روش انتقال میانگین در فیلتر ذره ای

در این بخش برای مقایسه الگوریتم پیشنهادی در بخش ۴-۲ با الگوریتمهای *TSHT* [۶] ، دنبال کننده مبتنی بر فیلتر ذره ای [۵] و دنبال کننده مبتنی بر انتقال میانگین [۲] ، این چهار الگوریتم را بر روی رشته های ۵،۴ و ۶ (شکل ۵-۵ ، ج و د) پیاده سازی کردیم.

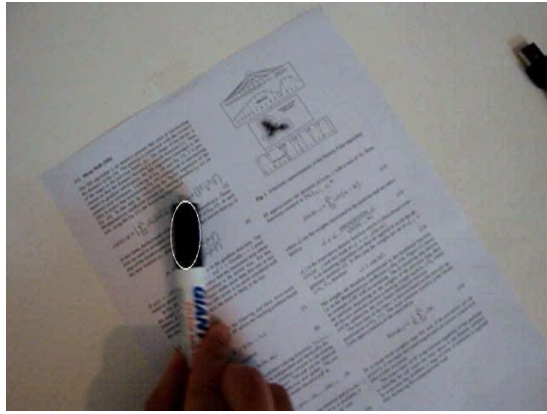
برای پیاده سازی دنبال کننده مبتنی بر فیلتر ذره ای ، الگوریتم ۳-۱ به کار گرفته شد. برای دنبال کننده *TSHT* نیز الگوریتم ۳-۱ با اضافه کردن یک مرحله اعمال انتقال میانگین بر روی تک تک ذره ها بعد از مرحله ۲ این الگوریتم پیاده سازی گردید و برای دنبال کننده پیشنهادی الگوریتم ۴-۱ پیاده سازی شد. دنبال کننده انتقال میانگین نیز بر طبق روش مرجع [۲] پیاده سازی شد.



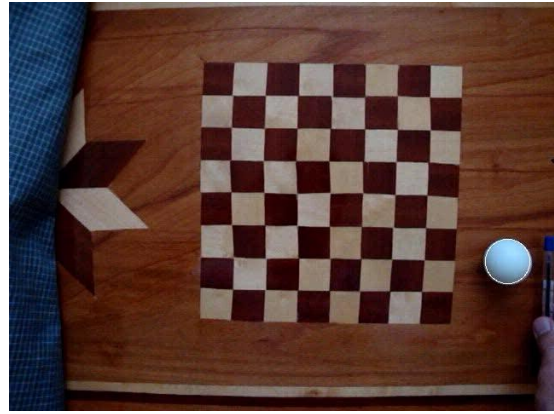
ب



الف



د



ج

شکل ۵-۶: فریم اول از الف) رشته ۳ که در آن شیء با دایره سبز رنگ مشخص شده است. ب) ، ج) و د) به ترتیب رشته های ۴، ۵ و ۶ که در آنها شیء با بیضی (یا دایره) سفید رنگ مشخص شده است.

پارامترهای به کار رفته در پیاده سازی این الگوریتمها مطابق بندهای زیر تعیین شدند :

الف) در هر چهار الگوریتم:

بیان شیء با استفاده از یک بیضی که ابعاد و موقعیت اولیه آن توسط کاربر تعیین می شود و با استفاده از هیستوگرام رنگ در فضای **RGB** با $16 \times 16 \times 16$ خانه صورت گرفته است.

ب) در الگوریتمهای **TSHT**، فیلتر ذره ای و الگوریتم پیشنهادی:

از معادله زیر به عنوان معادله پروسه استفاده شده است.

$$\mathbf{x}_k = \mathbf{x}_{k-1} + C_k \mathbf{v}_k \quad (1-5)$$

که در آن $\mathbf{x} = (x, y)$ و C_k به صورت زیر تعریف می شود.

$$C_k = \begin{bmatrix} c_{xk} & 0 \\ 0 & c_{yk} \end{bmatrix} = \begin{bmatrix} \frac{a}{r} \sum_{n=k-r}^k |x_n - x_{n-1}| & 0 \\ 0 & \frac{a}{r} \sum_{n=k-r}^k |y_n - y_{n-1}| \end{bmatrix} \quad (2-5)$$

و $\mathbf{v}_t = N(0, I)$ بیانگر بردار تصادفی با توزیع نرمال با میانگین صفر و ماتریس کواریانس یکانی می باشد. در

این رابطه ۲ $\mathbf{r} =$ و مقادیر اولیه c_{x1} و c_{y1} پنج انتخاب شد.

و وزن هر ذره را با استفاده از معیار مشابهت مربوط به رابطه (۳-۲۲) به شکل رابطه (۳-۵) تعیین کردیم.

$$\pi_k^{(n)} = \sum_{u=1}^m \sqrt{\hat{p}_u(s_k^{(n)})} \hat{q}_u \quad (3-5)$$

(ج) در الگوریتمهای $TSHT$ و الگوریتم پیشنهادی:

تعداد ذره ها (N_s)، ۲۰ عدد و $\mathbf{a}$ در رابطه (۳-۵) عدد ۱ انتخاب شد.

(د) در الگوریتم فیلتر ذره ای:

برای تعیین مناسب N_s و $\mathbf{a}$ به گونه ای که در وهله اول خطا و در وهله دوم زمان لازم برای اجرای الگوریتم

کاهش یابد بارها آزمایش صورت گرفت و در نهایت برای رشته ۴، $N_s = 30$ و برای رشته های ۵ و ۶ $= 100$

N_s و برای هر سه رشته $\mathbf{a} = 5$ انتخاب شدند.

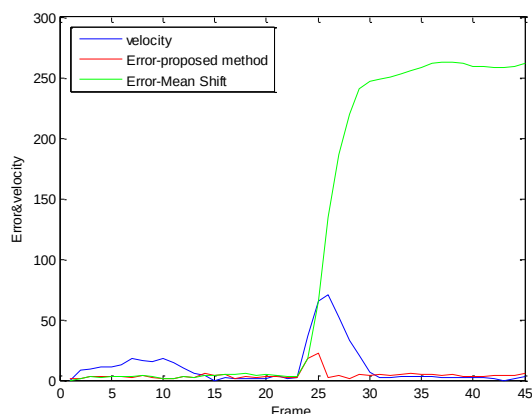
(ه) در الگوریتمهای $TSHT$ ، انتقال میانگین و الگوریتم پیشنهادی:

در هر سه روش، الگوریتم انتقال میانگین یکسان بر اساس مرجع [۲] و با تنظیمات مشابه به صورت زیر پیاده سازی شد:

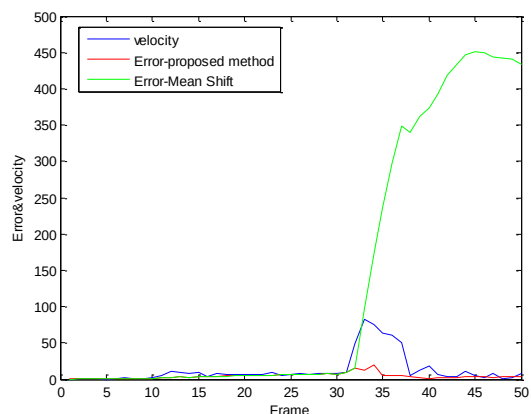
کرنل اپانچنیکف استفاده شد و شرط خاتمه انتقال میانگین رسیدن به ۲۰ تکرار و یا کوچکتر شدن مقدار جابجایی در دو تکرار پشت سر هم از مقدار آستانه $\varepsilon = 2$ تعیین شد.

برای رشته های استفاده شده موقعیت شیء در هر فریم به صورت دستی محاسبه شد و به عنوان مرجعی برای تعیین خطا در روشهای دنبال کردن مورد استفاده قرار گرفت.

در شکل ۵-۷ برای رشته های ۶ و ۵ خطای روشهای انتقال میانگین و دنبال کننده پیشنهادی در کنار سرعت شیء نمایش داده شده است. این نمودارها نشان می دهند که پس از افزایش سرعت جسم خطای دنبال کننده انتقال میانگین افزایش زیادی می یابد که نشان دهنده این مسئله است که دنبال کننده شیء را گم کرده است.



ب



الف

در جدول ۵-۱ نتایج مقایسه روش پیشنهادی با سه روش دیگر یعنی دنبال کننده TSHT ، دنبال کننده مبتنی بر فیلتر ذره ای و دنبال کننده مبتنی بر انتقال میانگین، آورده شده است. با توجه به تصادفی بودن موقعیت ذره ها در هر مرحله اجرای دنبال کننده TSHT ، دنبال کننده مبتنی بر فیلتر ذره ای و دنبال کننده پیشنهادی، اجرای این سه الگوریتم ، ۵ بار تکرار شد و میانگین گیری نتایج صورت گرفت . در این جدول اعداد ، نسبت را در مقایسه با نتیجه مربوط به روش پیشنهادی نشان می دهند برای مثال در رشته ۴ خطای روش TSHT ، 0.98 برابر خطای روش پیشنهادی است و زمان لازم برای اجرای الگوریتم TSHT در هر فریم، $11/71$ برابر زمان لازم برای اجرای الگوریتم پیشنهادی در هر فریم می باشد. خطا مانند بخش قبلی به صورت فاصله بین محل واقعی شیء و موقعیتی که دنبال کننده به عنوان موقعیت شیء تعیین کرده است می باشد. نتایج جدول ۵-۱ با میانگین گیری از نتایج مربوط به فریمهای هر رشته بدست آمده اند.

در مورد نتایج انتقال میانگین در مورد رشته های ۵ و ۶ از آنجا که دنبال کننده انتقال میانگین در رشته ۵ از فریم ۳۳ و در رشته ۶ از فریم ۲۵ به بعد شیء را گم می کند (شکل ۵-۷) برای اینکه مقایسه از نظر زمان منصفانه باشد همچنین نتایج مقایسه خطاها، قابل استفاده باشد، مقایسه برای رشته ۵ و ۶ به ترتیب تا فریم ۳۲ و ۲۴ صورت گرفته است.

جدول ۵-۱ : مقایسه روشهای TSHT ، دنبال کننده مبتنی بر فیلتر ذره ای و دنبال کننده مبتنی بر انتقال میانگین با روش پیشنهادی . اعداد در این جدول بیانگر نسبت خطا (یا زمان لازم برای اجرای الگوریتم) دنبال کننده مربوطه نسبت به خطا (یا زمان لازم برای اجرای الگوریتم) دنبال کننده پیشنهادی می باشند.

رشته ۴		رشته ۵		رشته ۶		میانگین	
خطا	زمان	خطا	زمان	خطا	زمان	خطا	زمان
۰/۹۸	۱۱/۷۱	۱/۰۱	۱۱/۹۸	۰/۸۷	۵/۹۸	۰/۹۵	۹/۸۹
۱/۸۸	۲/۲۱	۱/۶۱	۷/۳۴	۳/۰۲	۴/۲۲	۲/۱۷	۴/۵۹
۱/۰۴	۰/۹۴	تا فریم ۳۲: ۱	تا فریم ۳۲: ۰/۷۰	تا فریم ۲۴: ۱/۰۷	تا فریم ۲۴: ۰/۶۳	۱/۰۴	۰/۷۶

می دانیم یکی از نقاط قوت دنبال کننده مبتنی بر انتقال میانگین سرعت بالای آن است [۲] و یکی از نقاط ضعف آن این است که در صورتی که جابجایی شیء از یک فریم به فریم بعد از حدی بیشتر شود ، دنبال کننده در خطر گم کردن شیء قرار می گیرد.

در مواردی که دنبال کننده مبتنی بر انتقال میانگین در اثر سرعت بالای شیء آن را گم می کند می توان از الگوریتمهایی مانند فیلتر ذره ای و TSHT استفاده کرد . یک نقطه ضعف این الگوریتمها در مقایسه با دنبال کننده انتقال میانگین ، کند بودن آنهاست.

در اینجا می خواهیم این مسئله را نشان دهیم که در حالت کلی در مواردی که سرعت حرکت جسم خیلی کمتر از آن است که دنبال کننده مبتنی بر انتقال میانگین، جسم را به دلیل جابجایی زیاد آن گم کند ، سرعت الگوریتم پیشنهادی زیاد و نزدیک به سرعت الگوریتم دنبال کننده مبتنی بر انتقال میانگین خواهد بود. این مسئله بر خلاف الگوریتم انتشار چگالی شرطی [۵] است که در آن سرعت جسم تاثیری بر سرعت اجرای الگوریتم ندارد . در TSHT نیز بدون توجه به سرعت شیء به هر حال عملیات انتقال میانگین بر روی تمام ذره ها انجام میگیرد که این کار هزینه محاسباتی زیادی دارد.

در ادامه ابتدا به عوامل موثر بر احتمال ناکامی دنبال کننده مبتنی بر انتقال میانگین بر اثر سرعت بالای شیء می پردازیم و سپس بیان می کنیم که چگونه الگوریتم پیشنهادی در حالت کلی خود را با توجه به این عوامل وفق می دهد.

همانطور که قبلا گفته شد حداقل حوزه جذب برای یک نقطه همگرایی در روش دنبال کردن مبتنی بر انتقال میانگین ناحیه ای به اندازه شیء و به مرکزیت نقطه همگرایی است که آن را حوزه قطعی جذب نامیدیم.

ابتدا با ذکر یک مثال تاثیر میزان جابجایی شیء از یک فریم به فریم بعد، یا به عبارت دیگر سرعت شیء را بر احتمال ناکامی دنبال کننده مبتنی بر انتقال میانگین بررسی می کنیم. شکل ۵-۸ را در نظر بگیرید در این شکل دواير خاکستری نشان دهنده موقعیت شیء در فریم $k-1$ ام (فریم قبلی) و دواير سیاه نشان دهنده موقعیت شیء در فریم k ام (فریم جاری) می باشد.

در شکل (۵-۸الف) با توجه با اینکه موقعیت مرکز دایره مربوط به شیء در فریم قبلی، که نقطه آغاز انتقال میانگین در فریم جاری است، در محدوده شیء در فریم جاری قرار دارد، انتقال میانگین درست عمل خواهد کرد ولی در شکل (۵-۸ب) با توجه به اینکه موقعیت شیء در فریم قبلی خارج از محدوده شیء در فریم جاری می باشد انتقال میانگین دارای عملکرد تضمین شده نمی باشد و ممکن است به نقطه ای غیر از نقطه صحیح، همگرا شود.



ب

الف

شکل ۵-۸ جابجایی یک شیء از یک فریم به فریم دیگر (الف) جابجایی کم (ب) جابجایی زیاد

خارج شدن یا نشدن نقطه آغازین انتقال میانگین از محدوده ای که عملکرد صحیح انتقال میانگین در آن محدوده تضمین شده است علاوه بر سرعت شیء به ابعاد شیء نیز بستگی دارد برای مثال در شکل ۵-۹ الف جابجایی شیء منجر به خارج شدن نقطه آغازین انتقال میانگین از حوزه قطعی جذب موقعیت شیء می شود در حالی که در شکل ۵-۹ ب با همان مقدار جابجایی به دلیل بزرگتر بودن ابعاد شیء این جابجایی منجر به خارج شدن نقطه آغازین از محدوده عملکرد تضمین شده نمی شود.



ب

الف

شکل ۵-۹ اشیائی با ابعاد مختلف و جابجایی های یکسان (الف) جسم کوچک (ب) جسم بزرگ

حال می خواهیم نشان دهیم سرعت الگوریتم پیشنهادی ما هم، البته در حالت کلی و به طور غیر دقیق به همین دو عامل یعنی سرعت جسم و ابعاد آن بستگی دارد به طوری که در جایی که خطر ناکامی دنبال کننده انتقال میانگین کم است، سرعت الگوریتم پیشنهادی زیاد و در جایی که این خطر زیاد می شود الگوریتم پیشنهادی کندتر خواهد بود.

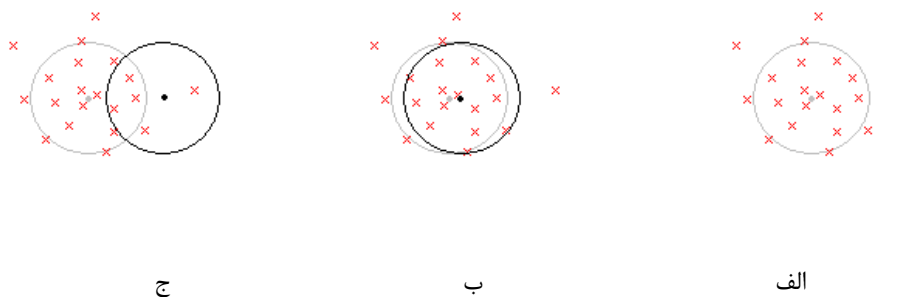
در اینجا به تاثیر این دو عامل یعنی سرعت شیء و ابعاد شیء بر سرعت الگوریتم پیشنهادی می پردازیم:

۱- ابعاد شیء: هر چقدر ابعاد شیء بزرگتر باشد، حوزه قطعی جذب مربوط به نقطه همگرایی هر ذره نیز بزرگتر خواهد بود بنابر این احتمال اینکه سایر ذره ها در آن قرار داشته باشند یا در طول فرایند اعمال انتقال میانگین بر روی آنها وارد این حوزه شوند بیشتر می شود بنابراین سرعت الگوریتم بالاتر خواهد بود.

۲- سرعت جابجایی شیء: در فریم k ام پس از اجرای مرحله ۲ الگوریتم پیشنهادی (الگوریتم ۴-۱) و قبل از آغاز مرحله انتقال میانگین، در صورت عملکرد صحیح الگوریتم، احتمالا تعداد زیادی از ذره ها در حوالی نقطه ای قرار خواهند گرفت که شیء در فریم $k-1$ ام در آن نقطه قرار داشته (شکل ۵-۱۰ الف) بنابراین اگر جابجایی موقعیت شیء در فریم k ام از موقعیت آن در فریم $k-1$ ام بسیار کم باشد می توان انتظار داشت حوزه قطعی جذب ذره ای که پس از انتقال میانگین به موقعیت صحیح شیء می رسد، تعداد زیادی از ذره های دیگر را در خود جای دهد و در نتیجه آن از اعمال انتقال میانگین بر روی تعداد زیادی از ذره ها بی نیاز شویم (شکل ۵-۱۰ ب و ج).

در شکل ۵-۱۰ نیز دواير خاکستري نشان دهنده موقعيت شىء در فریم $k-1$ ام و دواير سياه نشان دهنده

موقعيت شىء در فریم k ام می باشند. همچنین موقعيت هر ذره با علامت $\times$ نشان داده شده است.



شکل ۵-۱۰: تأثير میزان جابجايي شىء در سرعت الگوريتم (الف) توزيع ذره ها در فریم k ام پس از انجام مرحله دوم الگوريتم در مقايسه با موقعيت شىء در فریم $k-1$ ام (ب) جابجايي کم شىء (ج) جابجايي زياد شىء

علاوه بر تأثيری که سرعت جسم بر سرعت الگوريتم پيشنهادی دارد و توسط شکل (۵-۱۰) نشان داده شد ،

سرعت جسم از طريق ديگری نیز می تواند بر سرعت الگوريتم پيشنهادی تأثير گذار باشد. اين مسئله به

اين خاطر است که C_k (در رابطه ۵-۱) به سرعت شىء وابسته می باشد. در صورتی که حرکت جسم در 2

فریم قبلی زياد بوده باشد، با توجه به رابطه ۵-۲ مقادير c_{xk} و/يا c_{yk} افزايش خواهد يافت و بنابر با توجه به

رابطه ۵-۱ ذره ها از همدیگر بيشتتر دور می شوند. وقتی ذره ها از يکديگر دورند هنگامی که يکی از ذره ها

پس از اعمال انتقال میانگين بر روی آن به نقطه همگرایی خود می رسد ، احتمال اينکه ساير ذره ها در

حوزه قطعی جذب اين نقطه همگرایی قرار گیرند کمتر است . همچنین احتمالا ساير ذره هایی هم که در

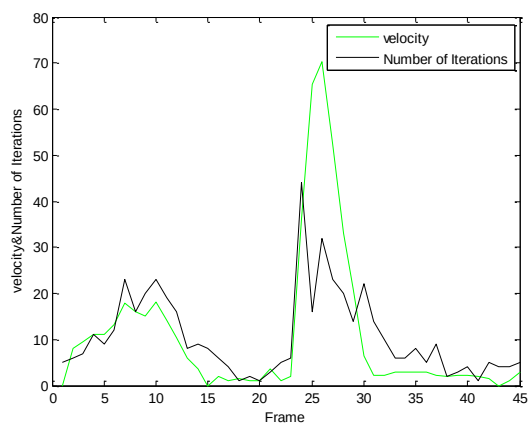
طول فرايند اعمال انتقال میانگين بر روی آنها به نقطه ای در اين حوزه می رسند بايد تا رسيدن به اين حوزه

مسافت بيشتري را طی کنند بنابراين الگوريتم سرعت کمتری خواهد داشت.

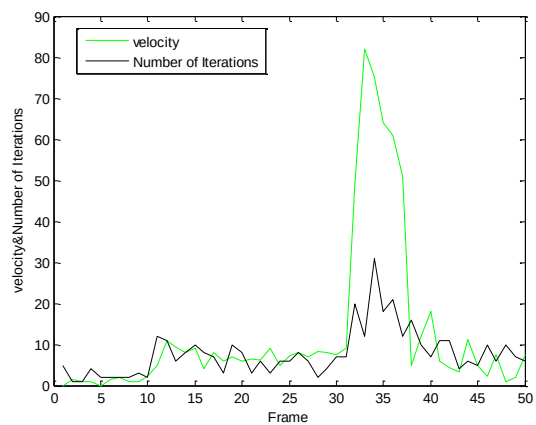
نمونه ای برای مطالب ذکر شده در این بحث نتایج آزمایشات مربوط به رشته ۴ می باشد که در این رشته سرعت جسم در مقایسه با ابعاد آن کم است . نتایج پیاده سازی الگوریتم پیشنهادی بر روی این رشته نشان می دهد در هر فریم ، با اعمال انتقال میانگین بر روی اولین ذره ، همه ذره های دیگر در حوزه قطعی جذب آن قرار می گیرند و بنابر این در هر فریم تنها یک بار انتقال میانگین اجرا می شود. و همانطور که در جدول (۵-۱) آمده زمان لازم برای اجرای الگوریتم پیشنهادی تنها ۶ در صد بیشتر از زمان لازم برای اجرای دنبال کننده مبتنی بر انتقال میانگین است.

در رشته های ۵ و ۶ هم می توان ارتباط سرعت حرکت شیء را با تعداد تکرارهای انتقال میانگین که در هر فریم اجرا می شود بررسی کرد . تعداد تکرارهای انتقال میانگین یک عامل اصلی در تعیین سرعت دنبال کننده پیشنهادی هستند به طوری که هر چقدر تعداد تکرارها بیشتر باشد، دنبال کننده به زمان بیشتری برای اجرا نیاز خواهد داشت و بالعکس.

در شکل (۵-۱۱) تعداد تکرارهای انتقال میانگین در هر فریم در مقایسه با سرعت جسم برای رشته های ۵ و ۶ آورده شده است. در این شکل پیروی غیر دقیق تعداد تکرارهای انتقال میانگین از سرعت جسم در هر فریم مشهود است. سرعت در این نمودارها برابر با میزان جابجایی شیء از فریم قبلی به فریم حاضر بر حسب تعداد پیکسل تعیین شده است.



ب



الف

شکل ۵-۱۱ تعداد تکرارهای انتقال میانگین در هر فریم در مقایسه با سرعت جسم الف) رشته ۵ ب) رشته ۶

فصل ششم

خلاصه و نتیجه گیری

یکی از روشهای محبوب برای دنبال کردن شیء در رشته ویدئویی، روش دنبال کردن شیء با استفاده از انتقال میانگین می باشد که در آن از شیوه انتقال میانگین برای پیدا کردن ناحیه ای در فریم جاری که بیشترین مشابهت را با تصویر مدل شیء دارد استفاده می شود. در دنبال کننده انتقال میانگین استفاده از هیستوگرام رنگ برای بیان شیء متداول است و این در حالی است که در مواردی که رنگهایی شبیه به رنگهای شیء در پس زمینه وجود دارد دنبال کننده ممکن است با مشکل مواجه شود. ما برای رفع این مشکل استفاده از اطلاعات جهت و دامنه گرادیان را در کنار اطلاعات رنگ پیشنهاد کردیم.

یکی از نقاط ضعف دنبال کننده انتقال میانگین این است که در صورتی که تغییر موقعیت شیء در یک فریم نسبت به فریم قبلی از میزان مشخصی بیشتر شود ممکن است دنبال کننده، شیء را گم کند. در مواردی که دنبال کننده انتقال میانگین به دلیل یاد شده در دنبال کردن شیء ناکام می ماند، می توان از دنبال کننده مبتنی بر فیلتر ذره ای استفاده کرد هرچند الگوریتم دنبال کردن مبتنی بر فیلتر ذره ای برای اجرا معمولاً نیاز به زمان بیشتری نسبت به الگوریتم دنبال کننده مبتنی بر انتقال میانگین دارد.

دنبال کننده دیگر مورد بررسی در این پایان نامه روش TSHT [۶] است که در آن یک مرحله شامل اعمال انتقال میانگین بر روی تک تک ذره ها به روش فیلتر ذره ای افزوده می شود. در این روش نیز چندین بار پیاده سازی عملیات انتقال میانگین، الگوریتم را کند می کند.

در بسیاری از کاربردها سرعت جسم به گونه ای است که دنبال کننده انتقال میانگین قادر به دنبال کردن آن می باشد . در مواردی ممکن است برای مدت بسیار کوتاهی در کل رشته ویدئویی، سرعت شیء تا حدی بالا باشد که باعث ناکام ماندن دنبال کننده انتقال میانگین شود. از طرفی الگوریتمهای فیلتر ذره ای و TSHT در این موارد قادر به دنبال کردن شیء هستند منتهی اجرای این الگوریتمها زمانبر می باشد. در این پایان نامه الگوریتمی را پیشنهاد نمودیم که خاصیت دنبال کنندگی سریع را مانند روش انتقال میانگین دارد در عین حال با افزایش سرعت شیء قادر به دنبال کردن آن خواهد بود.

در این دنبال کننده مانند روش TSHT یک مرحله شامل اعمال انتقال میانگین بر روی ذره ها به الگوریتم فیلتر ذره ای اضافه شده است با این تفاوت که در اینجا پس از اعمال انتقال میانگین بر روی یک ذره و بدست آوردن نقطه همگرایی ، ناحیه ای به اندازه شیء و مرکزیت آن نقطه همگرایی در نظر گرفته می شود و الگوریتم ، نقطه همگرایی سایر ذره ها در صورتی که در این محدوده قرار گیرند (یا در طول اعمال عملیات انتقال میانگین بر روی آنها وارد این محدوده شوند) را مرکز این حوزه در نظر می گیرد و از اعمال انتقال میانگین (و یا ادامه آن) بر روی آن ذره ها ، اجتناب می کند و به این وسیله سرعت اجرای الگوریتم افزایش می یابد.

نتایج عملی نشان دهنده سرعت قابل توجه این الگوریتم در مقایسه با دنبال کننده مبتنی بر فیلتر ذره ای و TSHT می باشد همچنین در آزمایشات مواردی نشان داده شده اند که دنبال کننده انتقال میانگین در اثر سرعت بالای شیء آن را گم میکند در حالی که دنبال کننده پیشنهادی شیء را دنبال می کند. با مقایسه

زمان لازم برای اجرای الگوریتم پیشنهادی و دنبال کننده مبتنی بر انتقال میانگین نشان دادیم که دنبال کننده پیشنهادی تنها کمی کندتر از دنبال کننده مبتنی بر انتقال میانگین عمل می کند.

کارهای آینده

۱- شکل بسیاری از اشیاء به گونه ای است که نمی توان آنها را با یک بیضی بیان نمود. تحقیق برای دستیابی به روشی برای بهبود در دنبال کننده مبتنی بر انتقال میانگین به گونه ای که بتوان در این دنبال کننده از روشهای دیگر برای بیان شیء استفاده کرد ، پیشنهاد می شود.

۲- یکی از مزایای دنبال کننده مبتنی بر انتقال میانگین در مقایسه با دنبال کننده های کانتور ، سرعت بالای آن می باشد .به عنوان کار آینده روشی را پیشنهاد می کنیم که در آن ، به کمک یک دنبال کننده انتقال میانگین ، موقعیت تقریبی شیء ، پیدا شود و پس از آن با استفاده از یک دنبال کننده کانتور ، شکل و موقعیت کانتور شیء تعیین گردد. انتظار می رود با این روش بتوان به دنبال کننده سریعتری در مقایسه با دنبال کننده های کانتور دست یافت به خصوص برای دنبال کردن اشیائی که سرعت زیادی دارند.

۳- می توان برای بیان شیء علاوه بر ویژگیهای رنگ ، دامنه گرادیان و جهت گرادیان از سایر ویژگی ها (مثلا ویژگی های بافت) نیز استفاده کرد. برای به کارگیری سایر ویژگی ها در بیان شیء می توانیم برای هر ویژگی جدید ، یک بعد به فضای هیستوگرام اضافه کنیم.

۴- در روش دنبال کننده مبتنی بر فیلتر ذره ای پس از وزن دهی ذره ها با میزان مشابهت آنها با مدل شیء ، الگوریتم باید موقعیت شیء در فریم جاری را تعیین کند. یک روش برای این کار تعیین موقعیت شیء در محل ذره ای که بیشترین وزن را دارد می باشد . و در روشی دیگر میانگین وزن دار شده موقعیت ذره ها به عنوان موقعیت شیء تعیین می شود. پیشنهاد ما در این زمینه این است که موقعیت وزن دار شده ذره ها را به عنوان داده ها در نظر بگیریم و با استفاده از شیوه انتقال میانگین ، محلی را که تراکم این داده ها در آن نقطه بیشترین مقدار است را جستجو کنیم و موقعیت این نقطه را به عنوان محل شیء در نظر بگیریم .این نقطه موقعیت ماکزیمم محلی تابع چگالی احتمال حضور شیء را نشان می دهد. در این روش می توان انتقال میانگین را از موقعیت ذره ای که بیشترین وزن را دارد آغاز کنیم.

فهرست مراجع

Mean Shift: A Robust Approach Toward Feature Space Analysis..1- COMANICIU, D. and MEER, P. 2002
.IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 603-619

- 2-COMANICIU, D., RAMESH, V., AND MEER, P. 2003. Kernel-based object tracking. *IEEE Trans. Patt. Analy. Mach.Intell.* 25, 564–575.
- 3- CUEVAS, E. , ZALDIVAR ,D. AND ROJAS ,R. 2005. Kalman filter for vision tracking. Technical Report B 05-12
- 4- ARULUMPALAM, M. S. , MASKEL ,S. , GORDON, N. AND CLAPP ,T. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. 2002 .*IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp.174–188.
- conditional density propagation for visual 5-ISARD , M. AND BLAKE ,A. CONDENSATION Computer Vision, 29(1) pp5-28. tracking.1998. *International. Journal of adaptive particle filter and Mean Shift tracker with*
- 6- MAGGIO, E. AND CAVALLARO,A. Hybrid Acoustics, Speech, and Signal Processing. transition model.2005. *Proc. Int. Conf.*
- 7-YILMAZ, A., JAVED, O., and SHAH, M. 2006. Object tracking: A survey. *ACM Comput. Surv.* 38, 4, Article 13(Dec. 2006), 45 pages. DOI = 10.1145/1177352.1177355
- 8-VEENMAN, C., REINDERS, M., AND BACKER, E. 2001. Resolving motion correspondence for densely moving points. *IEEE Trans. Patt. Analy. Mach. Intell.* 23, 1, 54–2.
- 9-SERBY, D., KOLLER-MEIER, S., AND GOOL, L. V. 2004. Probabilistic object tracking using multiple features. In *IEEE International Conference of Pattern Recognition (ICPR)*. 184–187.
- 10-YILMAZ, A., LI, X., AND SHAH, M. 2004. Contour based object tracking with occlusion handling in video acquired using mobile cameras. *IEEE Trans. Patt. Analy. Mach. Intell.* 26, 11, 1531–1536.
- 11-BALLARD, D. AND BROWN, C. 1982. *Computer Vision*. Prentice-Hall. BAR-SHALOM, Y. AND FOREMAN, T. 1988. *Tracking and Data Association*. Academic Press Inc.
- 12-ALI, A. AND AGGARWAL, J. 2001. Segmentation and recognition of continuous human activity. In *.IEEEWorkshop on Detection and Recognition of Events in Video*. 28–35
- 13-COOTES, T., EDWARDS, G., AND TAYLOR, C. 2001. Robust real-time periodic motion detection, analysis, and applications. *IEEE Trans. Patt. Analy. Mach. Intell.* 23, 6, 681–685.
- 14-ZHU, S. AND YUILLE, A. 1996. Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. Patt. Analy. Mach. Intell.* 18, 9, 884–900.
- 15-PARAGIOS, N. AND DERICHE, R. 2002. Geodesic active regions and level set methods for supervised texture segmentation. *Int. J. Comput. Vision* 46, 3, 223–247.

- 16-ELGAMMAL, A.,DURAIWAMI, R.,HARWOOD, D., ANDDAVIS, L. 2002. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. *Proceedings of IEEE* 90, 7, 1151–1163.
- 17-FIEGUTH, P. AND TERZOPOULOS, D. 1997. Color-based tracking of heads and other mobile objects at video frame rates. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 21–27.
- 18-EDWARDS, G., TAYLOR, C., AND COOTES, T. 1998. Interpreting face images using active appearance models. In *International Conference on Face and Gesture Recognition*. 300–305.
- 19-MUGHADAM, B. AND PENTLAND, A. 1997. Probabilistic visual learning for object representation. *IEEE Trans.Patt. Analy. Mach. Intell.* 19, 7, 696–710.
- 20-BLACK, M. AND JEPSON, A. 1998. Eigenttracking: Robust matching and tracking of articulated .objects using a view-based representation. *Int. J. Comput. Vision* 26, 1, 63–84
- 21-AVIDAN, S. 2001. Support vector tracking. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 184–191.
- 22-PARK, S. AND AGGARWAL, J. K. 2004. A hierarchical bayesian network for event recognition of human actions and interactions. *Multimed. Syst.* 10, 2, 164–179.
- 23-PASCHOS, G. 2001. Perceptually uniform color spaces for color texture analysis: an empirical evaluation.*IEEE Trans. Image Process.* 10, 932–937.
- 24-SONG, K. Y., KITTLER, J., AND PETROU, M. 1996. Defect detection in random color textures. *Israel Verj. Cap.J.* 14, 9, 667–683.
- 25-CANNY, J. 1986. A computational approach to edge detection. *IEEE Trans. Patt. Analy. Mach. Intell.* 8, 6,679–698.
- 26-HORN, B. AND SCHUNK, B. 1981. Determining optical flow. *Artific. Intell.* 17, 185–203.
- 27-LUCAS, B. D. AND KANADE., T. 1981. An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence*.
- 28-BLACK, M. AND ANANDAN, P. 1996. The robust estimation of multiple motions: Parametric and .piecewisessmooth flow fields. *Comput. Vision Image Understand.* 63, 1, 75–104
- 29-SZELISKI, R. AND COUGHLAN, J. 1997. Spline-based image registration. *Int. J. Comput. Vision* 16, 1-3, 185–203.

- 30-BARRON, J., FLEET, D., AND BEAUCHEMIN, S. 1994. Performance of optical flow techniques. *Int. J. Comput.Vision* 12, 43–77.
- 31-HARRIS, C. AND STEPHENS, M. 1988. A combined corner and edge detector. In *4th Alvey Vision Conference*.147–151.
- 32-SHI, J. AND TOMASI, C. 1994. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 593–600.
- 33-LOWE, D. 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60, 2,91–110.
- 34-SHI, J. AND MALIK, J. 2000. Normalized cuts and image segmentation. *IEEE Trans. Patt. Analy. Mach.Intell.* 22, 8, 888–905.
- 35-ROWLEY, H., BALUJA, S., ANDKANADE, T. 1998. Neural network-based face detection. *IEEE Trans. Patt. Analy.Mach. Intell.* 20, 1, 23–38.
- 36-VIOLA, P., JONES, M., AND SNOW, D. 2003. Detecting pedestrians using patterns of motion and appearance.In *IEEE International Conference on Computer Vision (ICCV)*. 734–741.
- 37-GREWE, L. AND KAK, A. 1995. Interactive learning of a multi-attribute hash table classifier for fast object recognition. *Comput. Vision Image Understand.* 61, 3, 387–416.
- 38-PAPAGEORGIOU, C., OREN, M., AND POGGIO, T. 1998. A general framework for object detection. In *IEEE International Conference on Computer Vision (ICCV)*. 555–562.
- 39-KUHN, H. 1955. The hungarian method for solving the assignment problem. *Naval Research Logistics Quart.* 2, 83–97.
- 40-TANIZAKI, H. 1987. Non-gaussian state-space modeling of nonstationary time series. *J. Amer. Statist. Assoc.*82, 1032–1063.
- 41-BROIDA, T. AND CHELLAPPA, R. 1986. Estimation of object motion parameters from noisy images. *IEEE Trans.Patt. Analy. Mach. Intell.* 8, 1, 90–99.
- 42-ROSALES, R. AND SCLAROFF, S. 1999. 3d trajectory recovery for tracking multiple objects and trajectory guided recognition of actions. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 117–123.
- 43-RANGARAJAN, K. AND SHAH,M. 1991. Establishing motion correspondence. *Conference Vision Graphies Image Process* 54, 1, 56–73.

- 44-WANG, J. AND ADELSON, E. 1994. Representing moving images with layers. *IEEE Image Process.* 3, 5, 625–638.
- 45-VIDAL, R. AND MA, Y. 2004. A unified algebraic approach to 2-d and 3-d motion segmentation. In *European Conference on Computer Vision (ECCV)*. 1–15.
- 46-BIRCHFIELD, S. 1998. Elliptical head tracking using intensity gradients and color histograms. In *IEEE .Conference on Computer Vision and Pattern Recognition (CVPR)*. 232–237
- 47-SCHWEITZER, H., BELL, J. W., AND WU, F. 2002. Very fast template matching. In *European Conference on Computer Vision (ECCV)*. 358–372.
- 48-BERTALMIO, M., SAPIRO, G., AND RANDALL, G. 2000. Morphing active contours. *IEEE Trans. Patt. . Intell.* 22, 7, 733–737.*Analy. Mach*
- 49-MANSOURI, A. 2002. Region tracking via level set pdes without motion computation. *IEEE Trans. Patt.Analy. Mach. Intell.* 24, 7, 947–961.
- 50-CREMERS, D. AND SCHNORR, C. 2003. Statistical shape knowledge in variational motion segmentation. *I. Srael Nent. Cap. J.* 21, 77–86.
- 51-RONFARD, R. 1994. Region based strategies for active contour models. *Int. J. Comput. Vision* 13, 2, 229–251.
- 52-BLAKE, A. AND ISARD,M. 2000. Active Contours:The Application ofTechniques from .*Graphics,Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion*. Springer
- 53-BADDELEY, A. 1992. Errors in binary images and an l version of the haus- dorff metric. *Nieuw Archief voor Wiskunde* 10, 157–183.
- 54-MACCORMICK, J. AND BLAKE, A. 2000. Probabilistic exclusion and partitioned sampling for multiple object tracking. *Int. J. Comput. Vision* 39, 1, 57–71.
- 55-BEYMER, D. AND KONOLIGE, K. 1999. Real-time tracking of multiple people using continuous ..detection. In *IEEE International Conference on Computer Vision (ICCV) Frame-Rate Workshop*
- 56-ISARD, M. AND MACCORMICK, J. 2001. Bramble: A bayesian multiple-blob tracker. In *IEEE International Conference on Computer Vision (ICCV)*. 34–41.
- 57-DOCKSTADER, S. AND TEKALP, A. M. 2001a. Multiple camera tracking of interacting and occluded human motion. *Proceedings of the IEEE* 89, 1441–1455.

- 58-MITTAL, A. AND DAVIS, L. 2003. M2 tracker: A multiview approach to segmenting and tracking people in a cluttered scene. *Int. J. Comput. Vision* 51, 3, 189–203.
- 59-COLLINS, R., LIPTON, A., FUJIYOSHI, H., AND KANADE, T. 2001. Algorithms for cooperative multisensor surveillance. *Proceedings of IEEE* 89, 10, 1456–1477.
- 60-CAI, Q. AND AGGARWAL, J. 1999. Tracking human motion in structured environments using a distributed camera system. *IEEE Trans. Patt. Analy. Mach. Intell.* 2, 11, 1241–1247.
- 61-LEE, L., ROMANO, R., AND STEIN, G. 2000. Monitoring activities from multiple video streams: Establishing a common coordinate frame. *IEEE Trans. Patt. Recogn. Mach. Intell.* 22, 8 (Aug.), 758–768.
- 62-KHAN, S. AND SHAH, M. 2003. Consistent labeling of tracked objects in multiple cameras with overlapping fields of view. *IEEE Trans. Patt. Analy. Mach. Intell.* 25, 10, 1355–1360.
- 63-KANG, J., COHEN, I., AND MEDIONI, G. 2003. Continuous tracking within and across camera streams. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 267–272.
- 64-HUANG, T. AND RUSSELL, S. 1997. Object identification in a bayesian context. In *Proceedings of International Joint Conference on Artificial Intelligence*. 1276–1283.
- 65-KETTNAKER, V. AND ZABIH, R. 1999. Bayesian multi-camera surveillance. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 117–123.
- 66-JAVED, O., RASHEED, Z., SHAFIQUE, K., AND SHAH, M. 2003. Tracking across multiple cameras with disjoint views. In *IEEE International Conference on Computer Vision (ICCV)*. 952–957.
- 67- LOWE , D. G. Distinctive image features from scale-invariant keypoints. 2004. *Int'l Journal of Computer Vision*, 60(2):91–110.
- DURAI SWAMI, R. AND DAVIS, L. 2005 . Fast Multiple Object Tracking via a Hierarchical Particle Filter. *Computer Vision. ICCV 2005. Tenth IEEE International Conference on*.
- 68-YANG ,C. ,

Abstract

Object tracking is one of important fields in machine vision. The aim of object tracking is positioning the object or objects in frames of a video sequence. In some applications extracting other information of moving object such as the region of the object is also needed.

In this thesis we address two main problems in an object tracking system namely object representation and process of tracking.

One important topic in an object tracking system is way to represent a moving object. In this thesis a new method for representing object in mean shift tracker is proposed. The existing methods use the statistical color information to represent an object. In the proposed method we use image gradient information beside the color information to improve the performance of existing object trackers. We have shown in presence of similar colors in object image and background, color based tracker fails to track the interest object while the proposed method performs successfully.

In case that the object movement is very fast particle filter based tracker performs superior to mean shift algorithm. In the second part of the thesis we propose an object tracking algorithm based on particle filtering in which particles are partially moved based on mean shift algorithm. The result of experiment on tracking samples indicates the proposed method runs quicker and more accurate than trackers based on conventional particle filtering. Moreover comparison between the proposed method and the mean shift tracker demonstrates that when the interest object moves very fast the proposed tracker successfully tracks the object while the mean shift tracker fails. Whereas it keeps to some extent the desirable quality of mean shift tracker named high rate of execution.

Object tracking, mean shift, particle filter, orientation histogram



Shahrood University of Technology
Faculty of Electrical, Electronic & Robotic Engineering

Object tracking in video sequence using statistical descriptors

Amir Nabaei

Supervisors:

Dr. Alireza Ahmadyfard

Dr. Ali Soleimani

February 2009