

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی کامپیوتر و فناوری اطلاعات
گروه هوش مصنوعی

استخراج ویژگی برای ردیابی انسان در شبکه‌ای از دوربین‌ها با دیدهای مجزا در محیط‌های بسته

علی قنبری سرخی

اساتید راهنما:
دکتر حمید حسن‌پور

استاد مشاور:
دکتر مرتضی زاهدی

پایان نامه جهت اخذ درجه کارشناسی ارشد
دی ۱۳۹۱

دانشگاه صنعتی شاهرود

دانشکده: مهندسی کامپیوتر و فناوری اطلاعات

گروه: هوش مصنوعی

پایان نامه کارشناسی ارشد آقای علی قنبری سرخی

تحت عنوان: استخراج ویژگی برای ردیابی انسان در شبکه‌ای از دوربین‌ها با دید مجزا

در تاریخ ۱۳۹۱/۱۰/۱۱ توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد مورد ارزیابی و با درجه عالی مورد پذیرش قرار گرفت.

امضاء	اساتید مشاور	امضاء	اساتید راهنما
	آقای دکتر مرتضی زاهدی		آقای دکتر حمید حسن‌پور

امضاء	نماینده تحصیلات تکمیلی	امضاء	اساتید داور
	مهندس فرهادی		دکتر علی اکبر پویان
			دکتر علیرضا احمدی‌فرد

تقديم به

چشمان منظر مادرم

دستان خسته پدرم

تشکر و قدردانی

سپاس خدایم که آدم را به نعمت تفکر آراست و اساتید فرزانه را
چون دکتر حمید حسینیور، دکتر مرتضی زاهدی و دکتر علی اکبر پویان را
در مسیر راهم قرار داد تا از اندیشه نابشان بهره گیرم و دانش و
بینششان را ره توشه خویشتن سازم. پاسرم دارم اندیشه بلندشان را و ارج
مرنهم هست و الایتان را.

تشکر مینمایم از پدر و مادر یگانه ام، برادر و خواهرانم که وجودشان
تکیه گاه برار تمام لحظه هار سخت من و دعاهایشان تنها سرمایه بال
گشودنم بسور خوشبختی است.

در نهایت بر خود لازم میدانم که از کمک هار جناب دکتر دثوف
غلام از دانشگاه کریمیه استرالیا و دکتر پیر لویجر مازیو از انیستیتور
تحقیق ایالتی کمال تشکر را داشته باشم.

تعهد نامه

اینجانب **علی قنبری سرخی** دانشجوی دوره کارشناسی ارشد رشته مهندسی کامپیوتر- گرایش هوش مصنوعی دانشکده مهندسی کامپیوتر دانشگاه صنعتی شاهرود نویسنده پایان نامه **استخراج ویژگی برای ردیابی انسان در شبکه‌ای از دوربین‌ها با دیدهای مجزا در محیط بسته** تحت راهنمایی آقای دکتر **حمید حسن پور** متعهد می‌شوم:

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

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها 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 بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

تحلیل و بررسی حرکات و رفتار انسان در حوزه‌های مختلف علمی همواره مورد توجه بسیاری از محققین بوده است. هدف از طراحی چنین سیستم‌هایی درک رفتار انسان از طریق توالی تصاویر ویدئویی است. تحلیل حرکات انسان از جنبه‌های متفاوتی از قبیل مقداردهی اولیه توالی تصاویر، ردیابی، بازیابی و شناسایی حرکات انسان قابل بررسی می‌باشد. ردیابی به عنوان یکی از مهمترین مراحل در مبحث تحلیل رفتار است که عموماً از طریق شبکه‌ای از دوربین‌ها که ناحیه وسیعی را تحت پوشش قرار می‌دهند، انجام می‌شود. جهت کاهش هزینه در چنین مواردی از دوربین‌هایی با دید مجزا استفاده می‌شود. اما ویژگی‌هایی مانند ظاهر افراد، شرایط نوری محیط، موقعیت افراد نسبت به دوربین و زاویه‌ی دید نتایج حاصله از چنین شبکه‌هایی را با خطا همراه می‌سازد.

در این پایان‌نامه راهکاری جهت ردیابی افراد در شبکه‌ای از دوربین‌ها با دید گسسته ارائه شده است. فرآیند ردیابی طی دو مرحله شناسایی و ردیابی افراد در درون ناحیه دید یک دوربین و انطباق بین افراد در سایر دوربین‌ها انجام گرفته است. پس از اعمال پیش‌پردازش‌های لازم و حذف قسمت‌های زائد سه ویژگی کلی مبتنی بر حرکت، مدل هندسی و ظاهر افراد مورد بررسی قرار گرفته و نقاط قوت و ضعف هریک به تفصیل ارائه شده‌است. با توجه به عدم ارائه نتایج مناسب توسط روش‌های مبتنی بر حرکت و مدل هندسی از هیستوگرام رنگ افراد به عنوان مهمترین ویژگی استخراج شده از ظاهر افراد استفاده شده است. هیستوگرام رنگ به دلیل سادگی و سرعت بالا جهت استفاده در این نوع از سیستم‌ها بسیار مناسب می‌باشد اما در محیط‌هایی با شدت روشنایی متفاوت و تغییر در اندازه افراد با مشکلاتی روبرو می‌شود. به همین منظور از تابع انتقال روشنایی تجمعی برای کاهش تغییرات روشنایی محیط در هنگام عبور افراد در بین دوربین‌ها استفاده شده است. جهت ردیابی با استفاده از ویژگی رنگ، بر اساس اندازه نسبی قسمت‌های مختلف بدن، بدن افراد به سه قسمت سر، نیم‌تنه و پایین‌تنه تقسیم شده‌است. این روش تحت تاثیر اندازه افراد قرار نمی‌گیرد و زمانی که افراد نسبت به

دوربین دور یا نزدیک می‌شوند بدن به درستی تقسیم‌بندی می‌شود. روش دیگری که در پایان‌نامه حاضر جهت استخراج ویژگی‌های لازم مورد استفاده قرار گرفته، روش تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی می‌باشد. ویژگی‌های معرفی شده توسط این روش را نمی‌توان همیشه در ورود و خروج افراد استفاده نمود زیرا هر یک از ویژگی‌ها در موقعیت‌های متفاوت محیطی نتایج متفاوتی را نشان می‌دهند. به همین منظور، با استفاده از سیستم فازی محیط تحت نظارت به شش قسمت فاصله نزدیک رو/پشت به دوربین، متوسط رو/پشت به دوربین و دور رو/پشت نسبت به دوربین تقسیم شد. به منظور ارزیابی عملکرد سیستم پیشنهادی آزمایشات بر روی مجموعه استخراج شده از آزمایشگاه‌های در کشور ایتالیا، پیاده‌سازی شده است. نتایج بدست آمده از روش پیشنهادی نشان دهنده عملکرد بهتر در محیط رنگی YCbCr نسبت به کارهای پیشین می‌باشد.

کلمات کلیدی: ردیابی افراد، شبکه‌ای از دوربین‌ها، زاویه دید مجزا، مدل ظاهری، هیستوگرام، فاصله نسبی بدن انسان، سیستم فازی، تابع انتقال روشنایی تجمعی، تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی.

لیست مقالات مستخرج از پایان نامه

۱. حمید حسن پور، علی قنبری سرخی، "ردیابی انسان در شبکه‌ای از دوربین‌ها با دید مجزا بر اساس هیستوگرام قسمت‌های مختلف بدن انسان، " مجله علمی پژوهشی رایانش نرم و فن آوری اطلاعات *Journal of Soft Computing and Information Technology (JSCIT)* ، دانشگاه صنعتی نوشیروانی بابل، ۱۳۹۱ (Accept)
۲. علی قنبری سرخی، حمید حسن پور، " ردیابی و شناسایی مجدد افراد در شبکه‌ای از دوربین‌ها با دید مجزا بر مبنای سیستم فازی در محیط‌های بسته، " نشریه مهندسی برق و الکترونیک ایران *Journal of Iranian Association of Electrical and Electronics Engineers*، دانشگاه امیرکبیر (Submitted)

فهرست

صفحه

عنوان

۱. مقدمه

۲	۲-۱- مقدمه‌ای بر ردیابی
۵	۲-۱- تعریف مسئله
۶	۳-۱- اهداف پایان نامه
۷	۴-۱- ساختار پایان نامه

۲. مروری بر عملکرد سیستم‌های ردیابی انسان

۱۰	۱-۲- مقدمه
۱۰	۱-۲- روش‌های نمایش شی
۱۲	۲-۲- ردیابی اشیاء
۱۵	۲-۲-۲- مدل‌های حرکت
۱۷	۲-۲-۲- مدل هندسی
۱۹	۳-۲-۲- مدل ظاهر
۲۴	۳-۲- شناسایی مجدد شی
۲۹	۴-۲- جمع‌بندی

۳. حذف پس‌زمینه و آشکارسازی هدف

۳۱	۱-۳- مقدمه
۳۱	۲-۳- روش‌های آشکارسازی و تشخیص هدف
۳۲	۱-۲-۳- روش تفاضل فریم‌های متوالی
۳۴	۲-۲-۳- روش حذف پس‌زمینه
۴۲	۳-۲-۳- روش جریان نوری
۴۳	۳-۳- حذف سایه
۴۵	۴-۳- روش مورد استفاده در پایان‌نامه به منظور آشکارسازی هدف
۵۰	۵-۳- جمع‌بندی

۴. ردیابی افراد در شبکه‌ای از دوربین‌ها دید مجزا

۵۲	 ۱-۴- مقدمه
۵۳	 ۲-۴- ردیابی و شناسایی مجدد افراد در چند دوربین
۵۴	 ۳-۴- بهبود عملکرد سیستم ردیابی افراد با روش پیشنهادی
۵۵	 ۱-۳-۴- موقعیت یا وضعیت افراد نسبت به دوربین
۵۹	 ۲-۳-۴- جدول وضعیت افراد
۶۲	 ۳-۳-۴- استفاده از روش فازی برای بدست آوردن وضعیت افراد نسبتاً به دوربین‌ها
۶۵	 ۴-۳-۴- غیر فازی کردن به روش ارتفاع
۶۶	 ۴-۴- استخراج ویژگی برای شناسایی افراد
۶۷	 ۱-۴-۴- رنگ
۶۸	 ۲-۴-۴- قسمت‌های مختلف بدن انسان
۶۹	 ۳-۴-۴- هیستوگرام قسمت‌های مختلف از بدن انسان
۷۲	 ۴-۴-۴- تابع انتقال روشنایی تجمعی
۷۴	 ۵-۴-۴- تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی
۷۸	 ۵-۴- فضای رنگی
۷۸	 ۶-۴- معیار شباهت
۸۱	 ۷-۴- ردیابی در یک دوربین با کمک روش پیشنهادی
۸۵	 ۸-۴- ردیابی در چند دوربین با زاویه دید مجزا با روش پیشنهادی
۸۶	 ۹-۴- جمع‌بندی

۵. پیاده‌سازی و ارزیابی نتایج

۸۹	 ۱-۵- مقدمه
۹۰	 ۲-۵- پایگاه داده
۹۴	 ۳-۵- بدست آوردن پارامترهای سیستم ردیاب در مرحله راه‌اندازی
۹۷	 ۴-۵- نتایج بدست آمده از وضعیت افراد نسبت به دوربین‌ها
۱۰۰	 ۵-۵- نتایج بدست آمده از تابع انتقال روشنایی تجمعی
۱۰۲	 ۶-۵- محاسبه تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی
۱۰۲	 ۷-۵- معیارهای ارزیابی
۱۰۳	 ۸-۵- نتایج بدست آمده برای ردیابی انسان در شبکه‌ای از دوربین‌ها

۹-۵- نتایج شناسایی مجدد افراد ۱۰۷

۱۰-۵- جمع بندی ۱۰۸

۶. نتیجه گیری و کارهای آتی

۱-۶- نتیجه گیری ۱۱۱

۲-۶- پیشنهادات ۱۱۵

فهرست اشکال

صفحه	عنوان
۵	شکل (۱-۱) تمام حالت‌های ممکن از خروج و ورود افراد در دو دوربین با دید گسسته
۱۲	شکل (۱-۲) انواع روش‌های نمایش شی
۲۵	شکل (۲-۲) ردیابی انسان در محیط با چند دوربین با دید مجزا
۲۵	شکل (۳-۲) توپولوژی کلی از سیستم‌های نظارتی با چند دوربین بدون هم‌پوشانی
۴۶	شکل (۱-۳) شمای کلی برای آشکارسازی اهداف در تصاویر ویدیویی
۴۶	شکل (۲-۳) تصویر زمینه برای دو دوربین با زاویه دید مجزا
	شکل (۳-۳) استخراج افراد از مجموعه مورد استفاده در این پایان‌نامه با استفاده از مدل
۴۹	گاوسی ترکیبی مخلوط تطبیقی
۵۴	شکل (۱-۴) شمایی کلی از مراحل ردیابی و شناسایی مجدد انسان
۵۶	شکل (۲-۴) نمایشی از ویژگی‌های طول، عرض و فاصله برای یک فرد
۵۷	شکل (۳-۴) نمایشی از افراد استخراج شده به همراه تصویر سایه‌نما
۵۹	شکل (۴-۴) بردار حرکت بدست آمده با روش Horn-Schunck
۶۳	شکل (۵-۴) مقدار هر ویژگی جدول (۲-۴) به عنوان ورودی به سیستم فازی
۶۴	شکل (۶-۴) وضعیت افراد نسبت به دوربین به عنوان خروجی سیستم فازی
۶۶	شکل (۷-۴) غیرفازی کننده به روش ارتفاع
۶۸	شکل (۸-۴) نسبت های اعضای و مدل مورد استفاده ی بدن انسان
۶۹	شکل (۹-۴) قسمت بندی بدن دو فرد با اندازه (قد) متفاوت با استفاده از روش پیشنهادی...
	شکل (۱۰-۴) هیستوگرام افراد متفاوت در فضای رنگی RGB و هیستوگرام هر قسمت از بدن
۷۱	به صورت جدا
۷۴	شکل (۱۱-۴) تاثیر شدت روشنایی و تاثیر تابع انتقال روشنایی بر تصاویر
۷۷	شکل (۱۲-۴) مراحل کلی الگوریتم تجزیه و تحلیل مولفه‌های اصلی قوی فازی
	شکل (۱۳-۴) مراحل کلی ردیابی و شناسایی مجدد افراد در شبکه‌ای از دوربین‌ها با دید
۸۱	مجزا.....
۸۴	شکل (۱۴-۴) شمای کلی از سیستم پیشنهاد به منظور ردیابی و شناسایی مجدد افراد
۸۶	شکل (۱۵-۴) ساختار سیستم ردیابی انسان برای شبکه‌ای از دوربین با دید مجزا
۹۰	شکل (۱-۵) نمای کلی از محیط تحت نظارت توسط دو دوربین با زاویه دید مجزا

- شکل (۲-۵) افراد موجود در مجموعه داده ۹۳
- شکل (۳-۵) دو نمونه از تصاویر مجموعه داده در دو دوربین متفاوت ۹۴
- شکل (۴-۵) نمونه‌ای از دو فرد با اندازه‌های متفاوت در دو دوربین ۹۴
- شکل (۵-۵) توابع فازی با استفاده از مقادیر جدول (۶-۵) برای دوربین شماره یک ۹۶
- شکل (۶-۵) توابع فازی با استفاده از مقادیر جدول (۶-۵) برای دوربین شماره دو ۹۷
- شکل (۷-۵) تابع انتقال روشنایی بدست آمده برای دو دوربین در مرحله راه‌اندازی ۹۷
- شکل (۸-۵) نتایج بدست آمده از روش فازی به منظور تعیین موقعیت افراد نسبت به دوربین‌ها ۹۹
- شکل (۹-۵) تاثیر تابع انتقال روشنایی در هنگام عبور افراد در بین دوربین‌ها ۱۰۱
- شکل (۱۰-۵) نمونه‌هایی از نتایج شناسایی و ردیابی افراد با روش‌های پیشنهادی در شبکه‌ای از دوربین‌ها ۱۰۷

فهرست جداول

عنوان	صفحه
جدول (۱-۳) مقایسه روش‌های آشکارسازی	۴۳
جدول (۱-۴) جدول وضعیت افراد نسبت به دوربین‌ها	۶۰
جدول (۲-۴) جدول وضعیت و مقادیر ویژگی‌های معرفی شده	۶۲
جدول (۳-۴) اندازه نسبی قسمت‌های مختلف	۶۹
جدول (۴-۴) میزان شباهت هیستوگرام افراد شکل (۴-۱۱) با استفاده از معیار شباهت	۷۲
جدول (۱-۵) افراد موجود در مجموعه داده که در زاویه دید دو دوربین در حال رفت و آمد هستند	۹۱
جدول (۲-۵) افراد موجود در مجموعه داده که در زاویه دید دوربین شماره یک در حال رفت و آمد هستند	۹۲
جدول (۳-۵) افراد موجود در مجموعه داده که در زاویه دید دوربین شماره دو در حال رفت و آمد هستند	۹۲
جدول (۴-۵) افراد موجود در مجموعه داده که در زاویه دید فقط یکی از دو دوربین در حال عبور می‌باشند	۹۳
جدول (۵-۵) مقادیر کمینه و بیشینه ویژگی‌های جدول وضعیت	۹۵
جدول (۶-۵) مقادیر بدست آمده برای جدول وضعیت در دو دوربین	۹۵
جدول (۷-۵) نتایج بدست آمده با استفاده از روش پیشنهادی برای شناسایی مجدد افراد در شبکه‌ای از دوربین‌ها با دید مجزا در محیط بسته	۱۰۷

فصل اول

مقدمه

۱-۱- مقدمه‌ای بر ردیابی

با گسترش فن‌آوری در دنیای امروز و نیاز روزافزون به ایجاد و توسعه سیستم‌های خودکار، نیاز به سیستم‌های هوشمند و مستقل از کاربر در بسیاری از کاربردهای مراقبتی، صنعتی، حمل و نقل، ذخیره و انتقال ویدیو دیجیتال و سیستم‌های ردیابی بیش از پیش به چشم می‌خورد. ردیابی افراد کاربردهای متفاوتی در حوزه‌های مختلف دارد که تمرکز این پایان‌نامه بر روی کاربردهای ردیابی افراد در سیستم‌های نظارتی محیط‌های هوشمند می‌باشد. از مهمترین کاربردهای این نوع از سیستم‌ها می‌توان به کنترل بر روی مکان‌های امنیتی، خانه‌های هوشمند، هوش تجاری ویدیویی^۱، مراقبت از افراد مسن و ناتوان و ... اشاره نمود. برای نمونه در کنترل مکان‌های امنیتی می‌توان به تشخیص رفتار-های مشکوک با استفاده از زیر نظر گرفتن حرکات فرد، دنبال کردن افراد در محیط و اعلام خطر هنگام ورود افراد به مناطق پرخطر و زیر نظر گرفت مجرمین اشاره کرد. یکی دیگر از کاربردهای که در بخش تجارت بسیار به آن توجه شده استخراج اطلاعات مفید برای بالا بردن میزان فروش و کسب رضایت مشتریان بوده، که در سال‌های اخیر توجه بسیاری از محققان و مدیران شرکت‌های بزرگ جهان را به خود جلب نموده است. ردیابی همچنین می‌تواند زمینه‌های برای زندگی بهتر و راحت‌تر را برای افراد مسن با ناتوانی‌های جسمی ایجاد نماید که به عنوان مثال برای مراقبت از این افراد و اتخاذ تصمیمات لازم در مواقع لزوم می‌توان اشاره نمود.

در ردیابی اشیاء متحرک، قبل از آنکه فرآیند ردیابی آغاز شود، ابتدا نیاز به تشخیص شیء متحرک یا جداسازی آن از پس‌زمینه می‌باشد. برای جداسازی شیء از پس‌زمینه مشکلاتی از قبیل وجود سایه، تغییر شدت روشنایی محیط، نویز و ... وجود دارد که شیء به خوبی از تصویر پس‌زمینه استخراج نمی‌شود برای این منظور باید پردازش‌های روی تصویر و پیش‌زمینه استخراج شده انجام داد تا کیفیت شیء استخراج شده بهبود پیدا کند.

¹ Video Business intelligence

بعد استخراج هدف که به عنوان ورودی سیستم ردیابی می‌باشد باید عمل تشخیص و ردیابی اهداف انجام شود. که در واقع باید تشخیص داد که شی که در فریم t در مکان x بوده در فریم $t+1$ در چه مکانی قرار دارد. ردیابی را می‌توان به عمل بدست آوردن، استنتاج و تخمین تغییرات زمانی-مکانی هدف یا روی هم رفته حالت‌های هدف در طول دنباله ویدئویی بر اساس اندازه‌گیری‌ها و مشاهدات تعریف کرد.

در سال‌های اخیر به منظور ردیابی از شبکه‌ای از دوربین‌ها استفاده شده است. استفاده از چند دوربین به دو دسته کلی شبکه‌ای از دوربین با دید مجزا یا مشترک تقسیم می‌شود که با توجه وسعت ناحیه دید و همچنین به صرفه بودن از لحاظ هزینه در این پایان‌نامه از شبکه‌ای از دوربین‌ها با دید مجزا استفاده شده است. یک مرحله مهم در این نوع از سیستم‌ها شناسایی مجدد افراد در زاویه دید دوربین‌ها می‌باشد در واقع تشخیص این که فردی که تازه وارد زاویه دید دوربین شده آیا در زاویه دید دوربین‌های دیگر دیده شده یا فرد جدید در سیستم می‌باشد.

بعضی از مفاهیم که در ردیابی بسیار کاربرد دارد به صورت زیر بیان می‌شود.

- شناسایی اشیاء: قطعه‌بندی تصویر به اشیاء و پس‌زمینه اولین و یک از مهمترین گام‌ها برای ردیابی انسان می‌باشد. شناسایی اشیاء (استخراج اشیاء) از پس‌زمینه تحت تاثیر عواملی مثل تغییر شدت روشنایی، شباهت رنگ اشیاء با پس‌زمینه، کیفیت پایین تصاویر، حرکت پس-زمینه مثل برگ درختان و سایه اشیاء کاری نسبتاً دشواری می‌باشد.
- ظاهر اشیاء: ظاهر اشیاء ممکن است بسته به شرایط و زاویه دید متفاوت باشد. مهم‌ترین دلیل آن می‌توان به دلیل تنوع در شکل جسم مخصوصاً حرکت از قسمت‌های مختلف بدن انسان باشد. به طور مشابه، زاویه دید و فاصله جسم نسبت به دوربین می‌تواند رایج‌ترین دلیل برای انطباق ناصحیح اشیاء از هم باشد. ظاهر شی در زمانی که مشاهدات از جلو، پشت و سمت‌ها دیگر باشد بسیار متفاوت است. به طور مشابه، ظاهر جسم نزدیک به یک دوربین ممکن است

بسیار متفاوت نسبت به زمانی که جسم فاصله زیادی دارد باشد. مشکلات انسداد یکی از مشکلات بزرگ در استفاده از ظاهر افراد می باشد چون زمانی که بخشی از قسمت های یک جسم قابل روئت نباشد ظاهر نمی تواند برای شناسایی افراد بسیار موثر باشد.

- شناسایی مجدد شی: زمانی که یک شی در زاویه دید یک دوربین قرار می گیرد سیستم نظارتی باید تصمیم بگیرد که این شی در مشاهدات گذشته (در این دوربین یا دوربین های دیگر) قبلا وجود داشته یا یک شی جدید می باشد. برای رسیدن به این هدف، نمایش اشیاء باید موقعیت و فاصله ثابت از دوربین باشد. اندازه و زاویه دید اشیاء زمانی که از یک دوربین به دوربین دیگر منتقل می شوند بسیار متفاوت می باشد.

ویژگی های استفاده شده به منظور ردیابی افراد عموما به سه دسته ویژگی های مبتنی به حرکت، مدل هندسی و ظاهر افراد تقسیم می شود.

ویژگی حرکت در ردیابی افراد در چند دوربین با دید مجزا نتایج رضایت بخشی از خود نشان نمی دهد زیرا حضور افراد در زمان و مکان های مختلف در مقابل دید دوربین ها تابع قانون خاصی نبوده و افراد می توانند به صورت غیر یکنواخت در زمان های مختلف در مکان های متفاوتی قرار گیرند و از آنجایی که ویژگی حرکت وابستگی زیادی به مکان و زمان افراد در فریم های قبلی دارد فیلترهای حرکت نمی توانند افراد را در مواقع انسداد و در زمان ورود به زاویه دید دوربین به خوبی شناسایی و ردیابی نمایند.

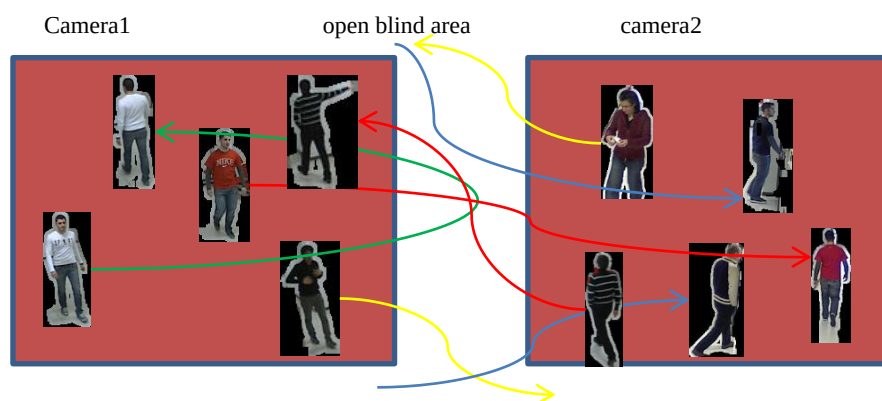
دسته دیگر از ویژگی ها که برای ردیابی افراد در شبکه ای از دوربین ها استفاده شده، مبتنی بر مدلی هندسی افراد می باشد، در این ویژگی، شناسایی لبه یک بخش اساسی را به خود اختصاص می دهد که به دلیل کیفیت پایین تصاویر ویدیوی نظارتی، نویز سنسورها و مشکلات قطعه بندی تصاویر، استخراج این ویژگی با مشکل روبرو می شود. مدل هندسی و ویژگی های آن در اشیاء غیر صلب به دلایلی مثل

اندازه کوچک، شکل متفاوت در توالی از ویدیو در قسمتی از اجسام متحرک، نتایج قابل قبولی از خود نشان نمی‌دهند.

ویژگی‌های مبتنی به ظاهر افراد در سال‌های اخیر بسیار برای ردیابی افراد در چندین دوربین با دید مجزا استفاده شده است.

۱-۲- تعریف مسئله

سیستمی با r دوربین $C_1, C_2, \dots, C_r$ داریم که در محیط بسته نصب شده‌اند. فضای دید این دوربین‌ها با هم هم‌پوشانی ندارد. فرض شده که n فرد $P_1, P_2, \dots, P_n$ در محیط در حال رفت و آمد می‌باشند. ردیابی در این نوع از سیستم‌ها طی دو مرحله شناسایی و ردیابی افراد در درون ناحیه دید یک دوربین، و انطباق بین افراد در دوربین‌ها صورت می‌گیرد. در این نوع سیستم، یک فرد ممکن است در زاویه دید دوربین، یک یا چندبار وارد یا خارج شود. شکل (۱-۱) تمام حالت‌های ممکن برای افراد در زمانی که از یک دوربین وارد/خارج می‌شوند را نشان می‌دهند. در نهایت در این پروژه می‌خواهیم سیستمی را طراحی نماییم که علاوه بر شناسایی انسان در یک دوربین، عمل ردیابی افراد در شبکه‌ای از دوربین‌ها انجام دهد.



شکل (۱-۱) تمام حالت‌های ممکن از ورود افراد در دو دوربین با دید گسسته

در این پایان‌نامه شرایط و فرضیاتی را برای محیط در نظر گرفته ایم که می‌توان به موارد زیر اشاره کرد:

الف) دوربین‌ها در مکان‌های ثابت نصب شده و مکان و زاویه دید آنها در زمان فیلم‌برداری تغییر نمی‌کند. با توجه به ثابت بودن مکان دوربین‌ها از مدل پس‌زمینه گاوسی برای حذف پس‌زمینه استفاده شده است.

ب) محیط در این پروژه بسته^۱ می‌باشد منظور از محیط بسته، ادارات و سازمان‌ها می‌باشد.

پ) شرایط نوری در محیط‌های مختلف متغیر ولی تغییرات آن محدود می‌باشد، برای از بین بردن تغییرات شرایط نوری در بین دوربین‌ها از یک تابع انتقال روشنایی استفاده شده است.

ت) فرض شده که ظاهر افراد در زمان خروج از یک دوربین و زمان ورود در این دوربین یا دوربین دیگر تا لحظه تشخیص، ثابت می‌باشد. در واقع افراد مجاز به تغییر لباس و نحوه پوشش خود در محدوده‌ی ورود و خروج از محیط نمی‌باشد. پس از این محدوده، فرد به شرطی که سایه‌نمای^۲ تشخیص داده شده توسط سیستم تغییر ننماید، می‌تواند نحوه پوشش خود را تغییر دهد.

ج) در این سیستم با توجه به اهمیت زمان پردازش کیفیت تصاویر پایین می‌باشد. وضوح بالای تصویر باعث بهبود کیفیت شناسایی و ردیابی سیستم ردیاب می‌شود ولی در این سیستم با توجه به اهمیت زمان پردازش کیفیت تصاویر به ابعاد ۳۲۰*۲۴۰ کاهش پیاده کرده است.

۳-۱- اهداف پایان نامه

با توجه به آنچه که بیان شد، هدف از این پایان نامه ارائه یک سیستم جامع به منظور ردیابی افراد در مجموعه‌ای از دوربین‌ها با دید مجزا (غیر هم‌پوشان) می‌باشد. جهت انجام این کار، فرآیند ردیابی به دو مرحله شناسایی و ردیابی افراد در زاویه دید یک دوربین و شناسایی مجدد افراد در زاویه بین دوربین‌ها تقسیم خواهد شد به منظور حذف پس‌زمینه از یکی از روش‌های آشکارسازی که دارای دقت و سرعت بالایی است استفاده خواهد شد. از آنجا که انتخاب ویژگی مناسب به منظور ردیابی در

^۱ Indoor

^۲ Silhouette

سیستم‌ها اهمیت بالایی دارد، در ادامه به معرفی تکنیک‌های ارائه شده پرداخته خواهد شد. در نهایت فرآیند کلی یک سیستم ردیاب در محیط‌های بسته به صورت مرحله به مرحله تشریح خواهد شد.

۱-۴- ساختار پایان‌نامه

این پایان‌نامه در شش فصل تهیه شده که هر فصل به شرح زیر می‌باشد.

فصل دوم- در این فصل به معرفی چگونگی عملکرد سیستم‌های ردیابی اشیاء و انسان پرداخته شده است. روش‌های موجود که در این زمینه در سال‌های اخیر توسط محققین انجام شده، معرفی گردیده است. در انتهای این فصل به معرفی سیستم‌های ردیابی با استفاده از شبکه‌ای از دوربین‌ها پرداخته شده است. این سیستم‌ها به دو دسته با زاویه دید مجزا و مشترک تقسیم شده‌اند که سیستم‌هایی با دید مجزا مورد بررسی قرار گرفته‌شده و به همین منظور مراحل و کارهای انجام شده به منظور شناسایی مجدد افراد به عنوان یک بخش بسیار مهم در شبکه‌ای از دوربین‌ها معرفی شده‌است.

فصل سوم- یکی از مهمترین مراحل در سیستم‌های ردیابی انسان، آشکارسازی هدف یا در واقع استخراج پیش‌زمینه از پس‌زمینه می‌باشد به همین منظور ما در این فصل به معرفی روش‌های موجود در حذف پس‌زمینه خواهیم پرداخت و نقاط قوت و ضعف هر روش را معرفی خواهیم نمود. در ادامه، به معرفی روش‌های بهبود عملکرد آشکارسازی هدف به مانند حذف نویز و سایه خواهیم پرداخت. در این پایان‌نامه با توجه به نقاط قوت روش مدل گاوسی مخلوط تطبیقی نسبت به روش‌های مطرح شده دیگر برای حذف پس‌زمینه استفاده شده و در انتهای این فصل نتایج مربوط به پیاده‌سازی با روش پیشنهادی برای طراحی سیستم ردیاب نمایش داده شده‌است.

فصل چهارم- به معرفی مراحل کلی روش شناسایی در یک دوربین و در مجموعه‌ای از دوربین‌ها با زاویه دید مجزا پرداخته شده است. در ابتدا روشی با استفاده از سیستم فازی برای تشخیص موقعیت افراد نسبت به دوربین‌ها معرفی شده و با توجه به اینکه یکی از مشکلات اساسی در شناسایی مجدد افراد تغییر شدت روشنایی می‌باشد به معرفی یکی روش قدرتمند در کاهش تغییر شدت روشنایی در

هنگام عبور افراد از زاویه دید دوربین‌ها پرداخته شده است. در قسمت‌های بعدی این فصل با توجه به نقاط قوت هیستوگرام این ویژگی به عنوان یکی از روش‌های سریع و موثر در ردیابی معرفی شده که برای افزایش قدرت این روش به تقسیم‌بندی قسمت‌های مختلف بدن انسان با استفاده از مطالعات پزشکی از نسبت‌های مختلف بدن انسان استفاده نموده‌ایم. در پایان یک روش دیگر با استفاده از تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی معرفی شده است.

فصل پنجم- پیاده‌سازی و ارزیابی نتایج: در این فصل ابتدا پایگاه داده استفاده‌شده و پیچیدگی‌های موجود در آن توصیف می‌گردد. سپس مراحل و پیاده‌سازی برای بدست آوردن موقعیت افراد نسبت به دوربین‌ها و تابع انتقال روشنایی در هنگام عبور افراد در بین دوربین‌ها توضیح داده شده است. در ادامه برای ردیابی و شناسایی مجدد از ویژگی‌های مطرح شده با توجه به موقعیت افراد نسبت به دوربین‌ها استفاده شده است. در پایان به معرفی روش‌های موجود برای ارزیابی سیستم مطرح پرداخته و نتایج بدست آمده از روش‌های پیشنهادی مورد ارزیابی قرار گرفته است.

فصل ششم- نتیجه‌گیری و کارهای آتی: جمع‌بندی و کارهای آتی در این فصل ذکر می‌گردد.

فصل دوم

مروری بر عملکرد سیستم های رویایی انسان

۲-۱- مقدمه

انتخاب یک روش مناسب برای نمایش شی می‌تواند در عملکرد ردیابی اشیاء تاثیر گذار باشد در همین راستا به معرفی روش‌های نمایش شی در تصاویر ویدیوی پرداخته شده است. در این فصل، به معرفی مطالعات در زمینه ردیابی انسان در یک دوربین و چند دوربین پرداخته شده است و مزایا و معایب روش‌های موجود شرح داده شده است. با توجه به اینکه استفاده از یک دوربین محدودیت وسعت دید را داشته به همین منظور از چندین دوربین برای ردیابی استفاده شده است. در بکارگیری شبکه‌ای از دوربین‌ها، دو دیدگاه کلی مطرح می‌شود که چند دوربین برای نظارت از یک صحنه استفاده شود یا دوربین‌ها به گونه‌ای در محیط نصب شوند که زاویه دید آنها هیچ هم‌پوشانی با هم نداشته باشد. در این پایان‌نامه با توجه به اینکه استفاده از شبکه‌ای از دوربین‌ها از لحاظ هزینه به صرفه بوده است این نوع سیستم بیشتر مورد توجه قرار گرفته است. شناسایی مجدد افراد یک مرحله مهم در این نوع از سیستم‌ها می‌باشد که در ابتدا به معرفی این مرحله پرداخته شده و سپس مطالعاتی که در این زمینه در سال‌های اخیر انجام شده، مورد بررسی قرار گرفته شده است. در نهایت با توجه به مطالبی که در این فصل ویژگی ظاهر افراد برای ردیابی در شبکه‌ای از دوربین‌ها با دید مجزا برای این پایان‌نامه انتخاب شده است.

۲-۲- روش‌های نمایش شی

در یک سیستم ردیابی، در ابتدا باید شی که قرار است ردیابی شود را معرفی کرد. اشیاء با شکل و ظاهر آنها بیان می‌شوند. نحوه نمایش می‌تواند در استخراج ویژگی از اشیاء بسیار حائز اهمیت باشد. هر یک از روش‌های نمایش شی با توجه به خصیصه شی که می‌تواند شی با اندازه کوچک/بزرگ یا صلب/ غیرصلب یا پیچیده یا باشد مورد استفاده قرار می‌گیرد. انتخاب روش مناسب برای نمایش شی می‌تواند عملکرد سیستم ردیابی را تحت تاثیر قرار دهد. از جمله روش‌های نمایش‌های شی که در الگوریتم‌های ردیابی متداول هستند را می‌توان به صورت زیر معرفی نمود.

الف) نقطه‌ها: شکل را می‌توان با یک نقطه یا مجموعه‌ای از نقاط نمایش داد [۱ و ۲]. در حالت کلی نمایش با نقطه برای ردیابی اشیائی که ناحیه‌ی کوچکی از تصویر را اشغال می‌کنند مناسب است. شکل (۱-۲-الف) و (۲-۱-ب) نمونه‌ای از این نوع روش‌ها را نشان می‌دهد.

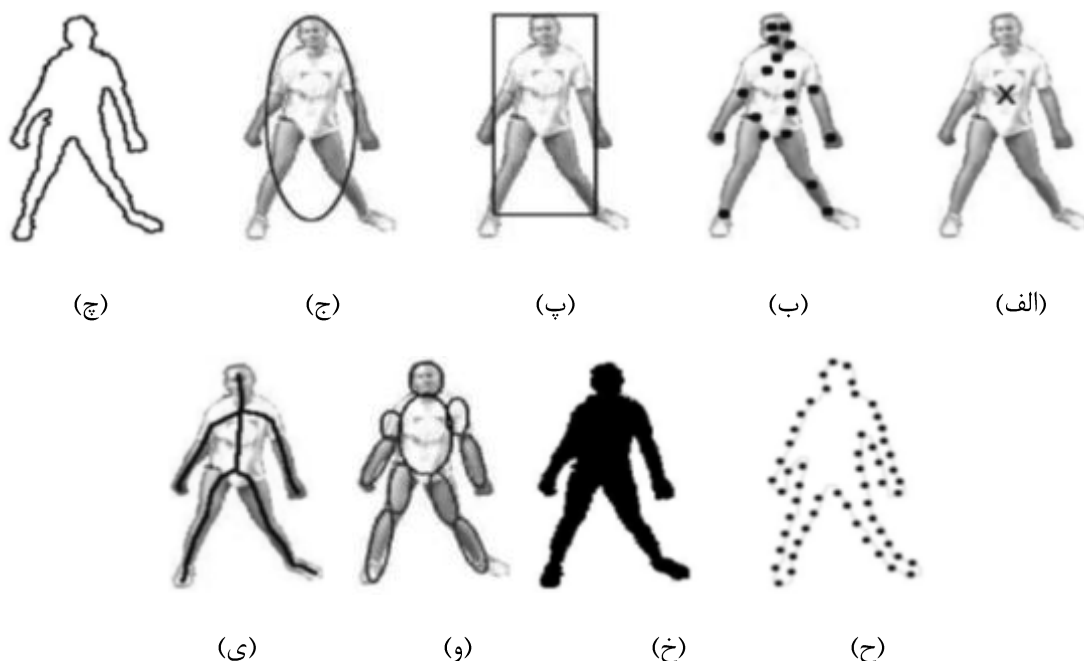
ب) اشکال هندسی اولیه (Primitive geometric shapes): شکل شی با یکی از اشکال هندسی پایه مانند مستطیل یا بیضی قابل نمایش است [۳] شکل (۲-۱-پ) و (۲-۱-ج) نمونه‌های از این روش را نشان می‌دهد. حرکت شی برای چنین نمایش‌هایی معمولاً با تبدیل‌های هندسی هوموگرافی، انتقال، مدل می‌شود. اگرچه اشکال هندسی اولیه برای نمایش اشیاء (rigid) غیر مفصلی مناسب‌تر هستند، لیکن برای ردیابی اشیاء مفصلی نیز به کار می‌روند.

پ) کانتور و سایه‌نمای شی (Object silhouette and contour): نمایش کانتور، مرز شی را تعریف می‌کند شکل‌های (۲-۱-چ) و (۲-۱-ح) نمایشی از آن را نشان می‌دهند. ناحیه‌ی درون مرز، سایه‌نمای شی نامیده می‌شود. نمایش کانتور و سایه‌نما برای ردیابی اشیاء پیچیده‌ی مفصلی مناسب هستند [۴]. شکل (۲-۱-خ) نمونه‌ای از نمایش سایه‌نمای شی می‌باشد.

ج) مدل‌های شکل بندبند (Articulated shape models): 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 [۵]. نمونه‌ای از این نمایش را در شکل (۲-۱-و) نشان داده است.

چ) مدل‌های اسکلتی (Skeletal models): در این نوع از نمایش شی از اسکلت بدن انسان استفاده می‌شود. این مدل نمایش را می‌توان برای مدل کردن اشیاء غیرمفصلی و همچنین اشیاء بندبند مورد استفاده قرار داد [۶]. نمونه‌ای از این نمایش را در شکل (۲-۱-ی) نشان داده است.

شکل (۲-۱) برخی از روش‌های ممکن برای نمایش شی را در تصاویر نشان می‌دهد. این ویژگی‌های شی برای تشخیص شی، مورد استفاده قرار می‌گیرند.



شکل (۱-۲) انواع روش‌های نمایش شی. (الف) مرکزی، (ب) مجموعه نقاط، (پ) مستطیلی شکل، (ج) بیضی شکل، (چ) مرز شی کامل (ح) نقاط مرکزی در مرز اشیاء، (خ) سایه نمای^۱ شی، (و) شکل بند، (ی) اسکلت شی [۵]

۲-۲- ردیابی اشیاء

ردیابی یک شیء در یک ویدئو به معنی شناسایی پیوسته‌ی موقعیت شی و به روز رسانی حرکت آن در زمانی است که هدف یا دوربین در حال حرکتند. با وجود اینکه کارهای زیادی در این زمینه انجام شده ولی ردیابی اشیاء به دلایل زیر هنوز برای محققین دارای اهمیت زیاد می‌باشد:

- از دست دادن اطلاعات، به دلیل اینکه یک دنیای سه بعدی توسط تصاویر دو بعدی مدل می‌شود.

- از دست دادن جزئیات در ویدئوهای نظارتی^۲ به دلیل پایین بودن وضوح دوربین‌ها و فشردگی زیاد فیلم

- ماهیت پیچیده حرکت و هندسه جسم

- طبیعت غیر صلب^۱ اشیاء مثل انسان

^۱ Silhouette

^۲ Surveillance

- هم‌پوشانی^۲ جزئی یا کلی شی با اشیاء دیگر و پس‌زمینه^۳
- تغییر در شرایط روشنایی^۴
- فرایند بلادرنگ^۵
- ماهیت غیرهمگن رنگ جسم در دوربین‌های متفاوت
- متفاوت بودن ظاهر اشیاء در محیط‌های با چند دوربین
- شباهت بین اشیا که می‌تواند به ارتباط نادرست اطلاعات منجر شود.

تحقیقات بسیار زیادی در جنبه‌های مختلف ردیابی اشیاء انجام شده که در این مطالعه به طور خلاصه به معرفی کارها و تکنیک‌های مرتبط که در این زمینه انجام شده می‌پردازیم.

در [۷] گزارش غنی از بررسی حدود ۴۰۰ مقاله منتشر شده‌است که به بررسی در زمینه مشکل ردیابی اشیاء و روش‌های اصلاح آن در حرکت انسان که شامل مدل اولیه انسان، ردیابی، تخمین مکان و شناسایی رفتار بوده است پرداخته‌اند و همچنین اشاره شده که مدل‌های کلی مورد نیاز است تا طیف گسترده‌ای از حرکات انسان را تحت پوشش قرار دهد.

در مطالعاتی که در [۵] انجام شده جزئیات بسیاری از تکنیک‌های موجود در ردیابی شی معرفی و گزارش خلاصه‌ای از این تکنیک‌ها بیان شده است. آنها روش‌های ردیابی را بر پایه ظاهر، شکل، رنگ و حرکت اشیاء دسته‌بندی کرده‌اند. در [۸] در مورد برخی از تکنیک‌های شناخته‌شده در ردیابی شی بحث کرده‌اند در بخش اول به بررسی این روش‌ها پرداخته و در بخش دوم، مقایسه عملکرد موجک^۶ [۹]، نمودار هیستوگرام از گرادیان جهت‌دار^۷ با SVM خطی [۱۰]، شبکه‌های عصبی [۱۱]، و ترکیبی

¹ Nonrigid

² Occlusion

³ Background

⁴ Illumination

⁵ Real-time

⁶ Wavelet

⁷ Histogram of Oriented Gradient (HOG)

از تشخیص با استفاده از شکل - بافت^۱ [۱۲] پرداخته‌اند. آنها ادعا می‌کنند بهترین نتایج را هنگامی که از ترکیب هیستوگرام با گرادیان جهت‌دار و SVM خطی استفاده شد از خود نشان داده است. این روش برای ردیابی چند شی در محیط بلادرنک به دلیل پیچیدگی هیستوگرام با گرادیان جهت‌دار و نیاز به پایگاه داده بزرگی از اشیاء در مرحله آموزش، مناسب نیست. در [۱۳] بررسی بخش‌های مختلف از شناسای انسان و ردیابی به طور مستقل و نه به طور کلی پیشنهاد کرده‌اند. آنها الگوریتم‌های ردیابی انسان را به پیش‌پردازش^۲، تقسیم‌بندی پیش‌زمینه، طبقه‌بندی شی، تطبیق^۳ / پالایش^۴، ردیابی و کاربردها تقسیم کردند. الگوریتم‌های مختلف در هر قسمت مورد بررسی قرار گرفت و نقاط قوت و محدودیت‌های آنها توضیح داده شده است. نحوه دسته‌بندی الگوریتم‌های ردیابی توسط بعضی از محققین به شیوه متفاوت انجام شده است. در این پایان‌نامه این روش‌ها به سه دسته با توجه به خواص ذاتی آنها تقسیم شده‌است. این دسته‌بندی‌ها بر پایه حرکت جسم، مدل‌های هندسی و ظاهر اشیاء می‌باشد. در بسیار از مواقع، بسیاری از ویژگی‌ها در دسته‌های متفاوت باهم ترکیب می‌شوند و نتایج بهتری را در شناسایی از خود نشان می‌دهند.

در ردیابی بر پایه حرکت، از پارامترهای حرکت مانند مکان، سرعت^۵ و شتاب بین افراد در فریم‌های پشت سر هم استفاده می‌شود. روش‌های فیلتر کالمن^۶، فیلتر ذرات^۷ و جریان‌های نوری^۸ از جمله معروف‌ترین روش‌ها در این دسته هستند.

مدل‌های هندسی از ویژگی‌های شکل برای شناسایی و ردیابی اشیاء استفاده می‌کنند. تشخیص لبه، تطبیق مرز شی، گشتاور^۱، مساحت، اندازه، از جمله ویژگی‌های شی محسوب می‌شوند که برای ردیابی مورد استفاده قرار می‌گیرند.

¹ Shape-texture

² Preprocessing

³ Verification

⁴ Refinement

⁵ Velocity

⁶ Kalman filter

⁷ Particle filter

⁸ Optical flow

تعداد زیادی از ویژگی‌ها و مدل‌ها برای شناسایی اشیاء و ردیابی وجود دارند اما مدل‌های برپایه ظاهر یک‌بعدی و دوبعدی بیشتر برای ردیابی رایج هستند. مدل مبتنی بر ظاهر به علت استحکام (قدرت بالا) تشخیص جسم صلب و غیر صلب محبوب‌تر هستند. در این روش مدل اشیاء فریم‌های قبلی با اشیاء فریم فعلی با استفاده از مدل ظاهر رنگ تطبیق داده می‌شود.

۲-۱-۲- مدل‌های حرکت

در ردیابی اشیاء بر پایه حرکت از مرکز جرم هندسی اشیاء استفاده می‌کند. به مرکز جرم اشیاء، مرکز ثقل^۲ گفته می‌شود و اگر تغییر کمی در شکل اشیاء به وجود بیاید تاثیر چندانی در مرکز ثقل به وجود نمی‌آید (مرکز جرم ثابت است). در ردیابی اشیاء براساس مدل حرکت جنبه‌های زیر را در نظر می‌گیریم:

- چگونگی حرکت اشیاء بین فریم‌ها
 - موقعیت زمانی و ظاهر شی در فریم‌های پی در پی
 - چگونگی تغییر سرعت اجسام متحرک
- حرکت شی با استفاده از فیلترهای خطی و گاهی اوقات غیرخطی مدل می‌شود. هر شی با استفاده از مدل حرکت ردیابی می‌شود. برچسب‌ها به اشیاء به طور مداوم بر اساس موقعیت واقعی و تخمین زده شده و سرعت، اختصاص داده می‌شوند. اگر چند شی در نزدیکی یکدیگر جابجا شوند، چنانچه ویژگی‌های حرکت تخمینی و واقعی آنها به هم نزدیک باشند، بعنوان یک شی در نظر گرفته می‌شوند.
- در ردیابی بر پایه حرکت به دلیل پیچیدگی این نوع از سیستم‌ها، فرض‌های را برای ردیابی در نظر می‌گیرند. به عنوان نمونه در [۱۴] فرض شده که که سرعت (نرخ) فریم‌های ویدئو به اندازه کافی بالا می‌باشد بنابراین موقعیت جسم به میزان قابل توجه‌ای بین فریم‌ها تغییر نمی‌کند. به طور مشابه [۱۵] فرض کردند که ظاهر جسم بین فریم و سرعت راه رفتن، متوقف شدن یا شروع به راه رفتن به

¹ Moments

² Center of gravity

تدریج تغییر می‌کند. همچنین در تحقیقات انجام شده در [۱۵ و ۱۶] و [۱۷] سرعت اشیاء را ثابت فرض شده است.

برخی از این فرضیات که در فوق به آنها اشاره شد می‌تواند صادق باشد. بدون در نظر گرفتن پارامترهای سناریوی خاص، که شامل موقعیت دوربین و تراکم جریان^۱ راه رفتن باشد. با این حال، این فرضیات در محیط‌های بدون محدودیت بی‌اعتبار است. به عنوان مثال، فرض سرعت ثابت به طور کلی برای عابر پیاده درست نیست. ممکن است تغییر مسیر و توقف افراد به صورت ناگهانی و غیرقابل پیش بینی انجام شود [۱۸].

در روش‌های ردیابی اشیاء برپایه حرکت نظیر فیلتر کالمن [۱۹]، مکان اشیاء در فریم بعدی با استفاده از تاریخچه حرکت اشیاء پیش‌بینی می‌شود. فیلتر کالمن مجموعه‌ای از رابطه‌های ریاضی برای پیش‌بینی به صورت بازگشتی را فراهم می‌کند در واقع حالتی از فرایند را به وسیله حداقل کردن متوسط مربع خطا تخمین می‌زند و پیش‌بینی حالت بعدی را با استفاده از این تخمین انجام می‌دهد. برای اطلاعات بیشتر در مورد جزئیات فیلتر کالمن و کاربردهای آن به [۲۰] مراجعه شود.

در [۲۱] از فیلتر کالمن برای محیط‌های توزیع شده با دوربین‌های بی‌سیم استفاده شده است. در هر دوربین موقعیت جسم با استفاده از فیلتر کالمن تخمین زده شده و موقعیت شی را به یک ایستگاه مرکزی فرستاده می‌شود. همچنین از فیلتر کالمن برای ردیابی شی و از مدل حرکت برای تشخیص اشیاء هم‌پوشانی در [۲۲] استفاده شده است.

فیلتر دیگری که برای ردیابی از حرکت استفاده می‌کند فیلتر ذره‌ای با دوام^۲ بوده است [۲۳]، که این روش با نمونه‌گیری آماری برای کاهش تعداد نمونه‌های مورد نیاز ترکیب می‌شود. مشکل اصلی در ردیابی انسان با فیلتر ذره‌ای رشد نمایی ذرات می‌باشد تا بتواند مسیر حرکت اشیاء را به درستی

^۱ Flow density

^۲ Annealed

تخمین بزند [۲۴ و ۲۵]. به همین دلیل این روش برای ردیابی چندین شی در کاربردهای بلادرنگ مناسب نیست.

در تحقیق دیگر، یک سیستم مجتمع با استفاده از مدل بدن انسان و حرکت در توالی از تصاویر در چندین دید متفاوت معرفی شده است [۲۶]، در این مطالعه از روش سلسله مراتبی مبتنی بر قواعد (پایه-نقش)^۱ برای مشخص کردن مکان و برچسب قسمت‌های بدن استفاده شده است. با توجه به دانش قبلی که از شکل قسمت‌های بدن، اندازه و ساختار آن برای قسمت‌بندی داشته است از یک فیلتر کالمن توسعه داده شده برای شناسایی حرکت انسان در بین فریم‌ها استفاده شده است. یک چارچوب^۲ ردیابی که از فیلتر کالمن استفاده کرده در [۲۷] پیشنهاد داده شده است.

به طور کلی، ویژگی‌های ساده حرکت برای ردیابی جسم به خصوص اگر اشیاء در حال حرکت انسان باشند مناسب نمی‌باشد زیرا حرکت انسان ممکن است به صورت تصادفی در فریم‌ها تغییر کند و فیلترهای حرکت قادر نباشند اشیاء را زمانی که باهم هم‌پوشانی دارند ردیابی نمایند. به خصوص اگر به طور طبیعی یک شی از زاویه دید یک دوربین خارج و بعد از مدتی دوباره وارد زاویه دید دوربین شود.

۲-۲-۲- مدل هندسی

مدل‌های هندسی از اشیاء و ویژگی‌های استخراج شده از آن در بسیاری از منابع برای شناسایی اشیاء مورد استفاده قرار گرفته است. در مدل هندسی به صورت گشتاور، لبه، اسکلت، مساحت، محیط و مدل شکل به صورت دوبعدی یا سه‌بعدی بیشتر مورد استفاده قرار گرفته است. البته بعضی از این ویژگی‌ها ممکن است وابسته به اندازه و یا فاصله جسم تا دوربین باشد اما الگوریتم‌های که نسبت به اندازه، انتقال و چرخش متغیر نمی‌باشند برای ردیابی و انطباق اشیاء می‌توانند استفاده شوند.

^۱ Rule-base

^۲ Framework

در بکارگیری ویژگی گشتاور جهت نمایش مدل هندسی یک جسم، از مجموعه نقاطی که هریک ویژگی خاصی از مدل هندسی را در خود دارد استفاده می‌شود. البته گشتاور ممکن نوع صفر، یک و ... باشد. گشتاور نوع صفر (مساحت اشیاء)، نوع یک (میانگین داده یا مرکز ثقل) و نوع دوم (پراکندگی داده) می‌باشند که برای آنالیز آماری در بسیار از موارد استفاده می‌شوند. برای انطباق اشکال اشیاء از چند نوع گشتاور بطور همزمان استفاده می‌شود. [۲۸] گشتاور ثابت که ترکیب خطی از گشتاور مرکزی هستند را معرفی کرد. [۲۹] شناسایی و ردیابی اشیاء را با استفاده برداری شامل سه گشتاور ثابت انجام داده است، و از فاصله اقلیدسی برای انطباق آنها استفاده کرده است. به طور مشابه [۳۰] از ویژگی‌های گشتاورهای اندازه، انتقال و چرخش با فرض انحراف نسبی^۱ اشیاء استفاده کرده‌اند. گشتاورهای در بعدهای دو، سه و n بعد در [۳۱] پیشنهاد شده است. در تحقیقات انجام شده، گشتاورهای دوبعدی نتایج خوبی در کاربردهای بینایی^۲ از خود نشان داده‌اند.

در [۳۲] ناحیه‌های مرزی بسته استخراج شد و نمایش این نواحی با استفاده از محیط، مساحت، فشردگی^۳ و گشتاورها پیشنهاد شده است. روش‌های مبتنی بر گشتاور به انحراف‌های اشیاء مانند غیر یکنواخت بودن شدت روشنایی و تغییر شکل اشیاء حساس می‌باشند. مرز شی معمولاً با تکنیک‌های شناسایی لبه استخراج می‌شود. در بعضی از تحقیقات، مرز شکل اشیاء از نقاط لبه پشت‌سرهم بدست آورده است [۳۳-۳۵]. برای پوشش نقاط مرزی از دیدگرایی که هر لبه را هم متصل می‌کند استفاده شده است [۳۶-۳۸]. بعضی از مواقع به دلیل کیفیت پایین تصاویر ویدیوی نظارتی، نویز سنسورها و مشکلات قطعه‌بندی تصاویر، شناسایی لبه با مشکل روبرو می‌شود.

¹ Affine distortion

² Vision application

³ Compactness

در مجموع مدل هندسی و ویژگی‌های آن در اشیاء غیرصلب به دلایلی مثل کیفیت پایین تصاویر، اندازه کوچک، شکل متفاوت در توالی از ویدیو در قسمتی از اجسام متحرک نتایج قابل قبولی از خود نشان نمی‌دهند.

۲-۲-۳- مدل ظاهر

در بسیاری از کارهای انجام شده از اطلاعات ظاهر در زمینه شناسایی و ردیابی افراد استفاده شده است. راه‌های زیادی به منظور استفاده از ظاهر برای شناسایی افراد در سال‌های اخیر معرفی شده است. از جمله آن می‌توان به رنگ، تابع چگالی احتمال، مدل‌های ظاهری یک‌بعدی و دوبعدی، چندین نقطه ثابت در اشیاء، شکل اشیاء و مدل ظاهری از چند زاویه دید، اشاره کرد.

تابع چگالی احتمال از ویژگی‌های ظاهری اشیاء (رنگ و بافت) می‌تواند، محاسبه شود. تابع چگالی احتمال یک شی می‌تواند به صورت پارامتری مثل گوسین [۳۹] و ترکیبی از گوسین‌ها^۱ [۴۰] و به صورت روش‌های غیر پارامتری برپایه کرنل [۴۱] و نمودار هیستوگرام [۴۲] باشند. نمودار هیستوگرام در دهه‌های اخیر بسیار مورد استفاده قرار گرفته است به دلیل اینکه نسبت به تغییر اندازه و چرخش مقاومت از خود نشان می‌دهد، هنوز هم یکی از روش‌های معروف برای شناسایی اشیاء می‌باشد. در [۴۲] از تکنیک جابجایی میانگین^۲ برای ردیابی اشیاء استفاده شده است. آنها از نمودار هیستوگرام رنگ و فاصله Bhattacharyya برای پیدا کردن بهترین انطباق موقعیت اشیاء استفاده کرده‌اند.

رنگ به عنوان مهمترین ویژگی برای شناسایی معرفی شده است. در [۴۳] از کانال‌های رنگی RGB استفاده شده و هر کانال در چهار برابر طول دامنه تعدیل داده شده^۳، و مکعب رنگ $4*4*4$ و هیستوگرام رنگ bin-۶۴ بدست آمده است. این تعدیل دادن تاثیر فضایی (مکانی) روشنایی رنگ متفاوت، را کاهش می‌دهد. از جمله کارهای که روش‌های ردیابی اشیاء به صورت خودکار (اتوماتیک)

¹ Mixture of Gaussians

² Mean shift technique

³ Quantize

نمی‌باشد می‌توان به [۴۴] اشاره کرد یعنی مرحله آموزش به صورت off-line می‌باشد. آنها تعدادی از bin‌های هیستوگرام را در فضای رنگی RGB انتخاب می‌کنند و مسیر بهینه اشیاء را با برنامه‌نویسی پویا بدست می‌آورند. [۴۵] از فشرده‌سازی بردار تردیجی^۱ برای تصاویر استفاده کردند و سپس با فاصله اقلیدسی میزان شباهت هیستوگرام بین این بردارها محاسبه کرده‌اند. مشکل اساسی تکنیک-های انطباق اشیاء برپایه هیستوگرام، از دست دادن اطلاعات فضایی (مکانی) سه بعدی^۲ می‌باشد. این مشکل در [۴۶] نیز ذکر شده است و در آن نشان داده‌شده که اگر برای شناسایی اشیاء از هیستوگرام بدون اطلاعات فضای سه‌بعدی استفاده کنیم عملکرد انطباق بین اشیاء کاهش می‌یابد.

تکنیک‌های غیرپارامتری مانند تخمین شدت کرنل برای مدل کردن ظاهر اشیاء مانند روش‌های قبلی بسیار محبوب هستند. [۴۱] ارتفاع اشیاء را به سه قسمت تقسیم می‌نمایند: سر^۳، نیم تنه^۴ و پایین تنه^۵ که هر قسمت با تابع شدت کرنل گوسین مدل شده‌است و این روش هزینه محاسبه تخمین شدت کرنل را بهبود داد. در این مقاله برای تقسیم قسمت‌های مختلف بدن، از تفاوت رنگ لباس پوشیده شده توسط هر فرد استفاده شده‌است. در واقع هر قسمت نیم‌تنه و پایین تنه باید کاملاً یک رنگ و نسبت به هم متفاوت باشد. [۴۷] از یک سیستم با چند دوربین با دید مشترک برای ردیابی انسان در محیط پیچیده استفاده کرده است. آنها ارتفاع اشیاء را به h برش^۶ (قسمت) و هر قسمت را با استفاده از تخمین شدت کرنل مدل کرده‌اند. و در ادامه این کار، در [۴۱] از n کرنل گوسین برای هر h برش استفاده شده که هزینه محاسبات را بالا برده است. [۴۸ و ۴۹] از تکنیک‌های برچسب‌گذاری بدن انسان استفاده کردند. اشیاء با استفاده از تئوری انطباق گراف شناسایی شد که این روش نتایج خوبی را از خود نشان می‌دهد ولی اگر اندازه اشیاء کوچک باشد عملکرد این روش کاهش پیدا می‌کند.

^۱ vector quantization

^۲ spatial

^۳ head

^۴ torso

^۵ Bottom part (legs)

^۶ slices

[۵۰] جنبه‌های بسیار زیاد ردیابی را مورد بررسی قرار داده‌است. آنها از اندازه افقی، مساحت، بافت افقی^۱، طرح افقی^۲ و شتاب حباب^۳ استفاده کردند. بافت افقی در [۵۰] یک مدل ظاهری یک‌بعدی است. این مدل ظاهری یک‌بعدی نه نرمال و نه تغییر اندازه در ارتفاع اشیاء را دارد و از ویژگی‌های هندسی مثل، طرح افقی و اندازه افقی برای انطباق بین اشیاء استفاده می‌کند. به همین دلیل این روش برای ردیابی انسان کمتر قابل اعتماد می‌باشد چون که انسان دارای اندازه‌های متفاوت در فریم‌های مختلف در چند دوربین با دید مشترک می‌باشد.

تکنیک‌های شناسایی اشیاء بر پایه چند نقطه^۴ از چند نقطه ثابت به عنوان ویژگی استفاده می‌کنند. این نقاط برای انطباق شیء فعلی با اشیاء موجود در پایگاه داده استفاده می‌شود. روش‌های زیادی برای انتخاب این نقاط ثابت (ویژگی‌ها) وجود دارد. [۵۱] تبدیل ویژگی‌ها با مقیاس ثابت (SIFT)^۵ را به عنوان روشی برای پیدا کردن این نقاط ویژگی معرفی کرده‌اند. [۵۲] یک الگوریتم سریع‌تر برای پیدا کردن نقاط کلیدی با عنوان ویژگی‌های قدرتمند با سرعت بالا (SURF)^۶ معرفی کرد. [۵۳] الگوریتم SIFT و SURF را مورد بررسی قرار دادند و ادعا کردند که این دو الگوریتم از لحاظ عملکرد شبیه همدیگر عمل می‌کنند اما SURF نسبت به چرخش و تغییر روشنایی ثابت نیست و از لحاظ سرعت محاسبات از SIFT سریع‌تر است. از نقاط ضعف این روش‌ها: الف) در ابتدا باید مرحله آموزش داشته باشیم و فقط اشیاء که در مرحله آموزش وجود دارند شناخته می‌شوند. ب) هر دو روش از لحاظ بار محاسباتی بسیار زمانبر هستند و برای کارهای بلادرنگ مناسب نمی‌باشند.

ایده انطباق الگو^۷ در واقع خلق یک مدل برای اشیاء می‌باشد. این الگو با یک تصویر یا حباب پیش-زمینه^۱ برای شناسایی انطباق داده می‌شود. در کل برای گرفتن یک الگو، مکان، مقیاس و جهت

¹ vertical texture

² horizontal projection

³ blob acceleration

⁴ Multi-point

⁵ Scale Invariant Feature Transform(SIFT)

⁶ Speed Up Robust Features(SURF)

⁷ template

استفاده می‌شود تا شباهت بین اشیاء (با در نظر گرفتن یک حدآستانه) اندازه‌گیری شود. الگوها از هندسه ساده اشکال یا سایه‌نما^۲ بدست می‌آیند [۵۴]. یک توسعه در الگوها، از اطلاعات فضایی (مکانی) و ظاهری باهم استفاده کرده است. از آنجایی که الگوها، ظاهر افراد را فقط از یک زاویه دید رمزگذاری^۳ می‌کنند بنابراین این روش برای ردیابی اشیاء که حالت آنها زیاد در طول ردیابی تغییر نمی‌کنند مناسب هستند. [۵۵] حدود ۱۱۰۰ الگو از طریقه راه رفتن نیم رخ که این تعداد الگو محاسبات بالا برای انطباق طریقه راه رفتن نیاز دارد را بدست آورند. مخصوصا اگر در هر الگو مکان‌ها، جهت و مقیاس متفاوت را جستجو کنیم. در نهایت لازم به ذکر است، اگرچه سلسله مراتب الگو^۴ می‌تواند تعداد متنوعی از اشکال اشیاء را تحت پوشش قرار دهد ولی این روش برای کنترل اشکال بزرگ تغییرپذیر وقتی که خیلی نزدیک به دوربین هستند مناسب نیستند [۵۶]. به همین دلیل این روش نمی‌تواند برای ردیابی که دوربین محیط‌های آنها تغییر قابل ملاحظه‌ای باهم دارند عملکرد خوبی از خود نشان دهد.

مدل‌های با چند-نما (نمایش)^۵ در واقع رمزگذاری یک شی با دیدهای متفاوت است. یک روش برای نمایش نماهای اشیاء متفاوت این است که یک زیرفضا^۶ (شبه فضا) از نماها بدست بیاوریم. برای نمونه از روش‌های زیرفضا می‌توان به تجزیه و تحلیل مولفه‌های اصلی (PCA)^۷ و تجزیه و تحلیل مولفه‌های مستقل (ICA)^۸ اشاره کرد. که برای نمایش ظاهر و شکل استفاده شده‌اند [۵۷ و ۵۸].

[۵۹] الگوریتم‌های ردیابی اشیاء برپایه یک دوربین ثابت (دوربین اصلی) و چندین دوربین متحرک دادند. آنها بحث‌های زیادی در مورد شناسایی ویژگی‌هایی به مانند، هیستوگرام رنگ، هیستوگرام از

¹ foreground blobs

² silhouettes

³ encode

⁴ template hierarchy

⁵ Multi-view

⁶ subspace

⁷ Principal Component Analysis (PCA)

⁸ Independent Component Analysis (ICA)

گرا دیان جهت‌دار، SIFT، SURF انجام دادند. [۶۰] حرکت و هیستوگرام رنگ را با هم ترکیب کرده و ادعا کرده که نتایج خوبی گرفته است.

به تازگی، بسیاری از محققان از چندین دوربین با دید مشترک برای ردیابی اشیاء که با هم هم‌پوشانی دارند استفاده کرده‌اند [۴۷ و ۶۱-۶۲]. مزایای محیط با چند دوربین با دید مشترک این است که چندین نما از یک اشیاء می‌دهد و برای انطباق اشیاء تحت برخورد (هم‌پوشانی) می‌توانیم از زاویه دید (نما) دوربین‌های دیگر استفاده نماییم. با وجود مزایای این دوربین‌های که می‌توان از پارمتر عمق به همراه کنترل هم‌پوشانی اشیاء استفاده کرد ولی مشکلاتی مثل کالیبره شدن دوربین‌ها با هم، تغییر در فضای دید یک دوربین موجب تغییر در بقیه دوربین‌ها می‌شود و از طرفی محدوده دید تحت پوشش نیز بسیار کم می‌باشد، اشاره کرد.

ردیابی برپایه ظاهری، بسیار برای سیستم‌های نظارتی مورد استفاده قرار گرفته است. این روش نتایج رضایت بخشی در شرایط متفاوت مثل کیفیت پایین تصاویر، بزرگ یا کوچک بودن اندازه اشیاء از خود نشان داده‌اند. این روش‌ها همچنین می‌توانند برای شناسایی مجدد^۱ اشیاء وقتی که از زاویه دید یک دوربین خارج شده و به زاویه دید دوربین دیگر وارد می‌شوند از خود نشان دهند.

۲-۳- شناسایی مجدد شی

کار مشاهده یک شی در زاویه دید یک دوربین و شناسایی دوباره همان شی در زاویه دید همان دوربین یا دوربین دیگر را شناسایی مجدد شی گویند. یک شی ممکن است در زاویه دید دوربین، یک یا چندبار وارد شود. اگر الگوریتم‌های بتوانند کار شناسایی مجدد یک شی و اطلاعات مربوط به زمان و تعداد مرتبه‌های ورود یا خروج این شی از زاویه دید یک دوربین را بدست بیاورد، این کار می‌توان کمک بسیار مهمی برای شناسای و آنالیز رفتار اشیاء انجام دهد. شناسایی مجدد اشیاء همچنین می-

¹ Re-Identification

تواند برای مراقبت (نظارت) از مکان‌های با وسعت بالا مثل: فرودگاه‌ها، محوطه دانشگاه، مراکز خرید و ایستگاه قطار با استفاده از چند دوربین بسیار مهم باشد.

در محیط‌های با چند دوربین، شناسایی مجدد شیء کمک می‌کند تا رکوردهای ثابتی از اشیاء متحرک وقتی که این اشیاء از زاویه دید دوربین خارج می‌شوند و دوباره وارد همان دوربین یا زاویه دوربین دیگر می‌شوند، بسازیم. بسیاری از محققان از دوربین‌ها با زاویه دید مشترک برای ردیابی استفاده کرده‌اند. که استفاده از این نوع محیط‌ها داری مزایای زیر می‌باشند:

اشیاء می‌توانند حتی زمانی که این اشیاء در یک یا چند دوربین باهم برخورد دارند، شناسایی شوند.

عملکرد انطباق اشیاء و شناسایی بهبود پیدا می‌کند

شکل و جزئیات سه‌بعدی می‌توان از اشیاء بدست آورد.

ولی ناحیه تحت پوشش این دوربین‌ها بسیار محدود می‌باشد. به همین منظور از محیط‌های با چند دوربین با دید مجزا (بدون هم پوشانی) برای نظارت ناحیه وسیعی استفاده می‌شود. در این بخش ما به بررسی، شناسایی مجدد اشیاء در چندین دوربین بدون هم‌پوشانی می‌پردازیم.



Cam1

Cam2

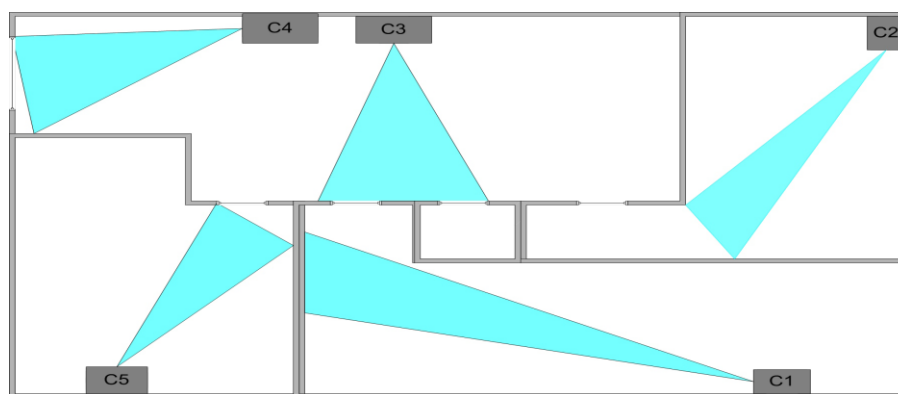
Cam3

شکل (۲-۲) ردیابی انسان در محیط با چند دوربین با دید مجزا [۶۳]

شکل (۲-۲) یک نمونه خاص از ردیابی انسان در محیط با چندین دوربین بدون هم‌پوشانی را نشان می‌دهد. در این محیط سه دوربین نصب شده که دو دوربین در بیرون و یک دوربین در داخل

ساختمان قرار دارد. به اشیاء یک برچسب وقتی که از یک دوربین خارج و در دوربین دیگر وارد می‌شوند اختصاص داده می‌شوند.

[۶۳] از پارامترهای فضایی (مکانی) و حرکت شی برای ردیابی در این نوع محیط‌ها استفاده کرده است. در این مقاله این پارامترها با مدل‌های ظاهری ترکیب شد و نشان داده شده که این کار سبب بالا رفتن عملکرد شناسایی مجدد افراد شده است. شکل (۲-۳) یک نمایش کلی از محیط‌ها با چندین دوربین با دید مجزا می‌باشد. در این شکل، پنج دوربین در یک ساختمان برای نظارت نصب شده‌است.



شکل (۲-۳) توپولوژی کلی از سیستم‌های نظارتی با چند دوربین بدون هم‌پوشانی [۶۴]

در این قسمت به بررسی روش‌های موجود برای شناسایی مجدد اشیاء در محیط‌ها با دید مجزا می‌پردازیم.

ویژگی حرکت نمی‌تواند برای دوباره‌شناسی شی در محیط‌های با دوربین‌های با دید مجزا مفید باشد زیرا نقاط ورود و خروج زیادی وجود دارند این در حالی است که ویژگی حرکت به مکان شی بستگی دارد. این امکان وجود دارد که اشیاء از یک مکان خارج شوند و دوباره در همان صحنه ولی این بار از مکان دیگر وارد شوند. به طور مشابه خواص هندسی نمی‌توانند نتایج خوبی برای شناسایی دوباره اشیاء برای اشیاء غیرصلب به دلیل تغییر اندازه بسیار زیاد (دور و نزدیک شدن شی نسبت به دوربین‌های متفاوت)، از خود نشان دهند.

برخی از محققان از ویژگی حرکت برای شناسایی مجدد اشیاء در چندین دوربین بدون هم‌پوشانی استفاده کرده‌اند آنها برای ردیابی اشیاء ویژگی‌های حرکت را با اطلاعات فضایی از دوربین‌ها ترکیب

کرده‌اند. [۶۵-۶۷] از اطلاعات بدست آمده از مشاهده مکان و سرعت اشیاء متحرک در میان چندین دوربین با دید مجزا برای تعیین ارتباط فضایی (زمان و مکان) بین دوربین‌ها استفاده کرده‌اند. [۶۶] فرض کرده که ارتباط بین اشیاء در بین این دوربین‌ها مشخص است. برای مدل کردن دینامیک افراد از یک مدل مارکف استفاده شده‌است. [۶۸] در این مقاله از ترکیب اطلاعات زمانی و مکانی به همراه ویژگی رنگ، تابع انتقال روشنایی برای برقراری ارتباط بین افراد استفاده شده است. این دسته از تحقیقات برای مواقعی مناسب می‌باشند که فاصله بین دوربین‌ها کم بوده و افراد ملزم به حفظ وضعیت حرکتی خود از جمله سرعت و جهت حرکت باشند. در صورتی که افراد بین دو دوربین مکث نموده و یا حرکت آزادانه‌ی دیگر انجام دهند سیستم دچار اختلال در تشخیص می‌گردد.

در بسیاری از مقالات منتشر شده از تکنیک‌های انطباق هیستوگرام شی مانند هیستوگرام رنگ به عنوان یک مشخصه^۱ استفاده شده است [۶۹]. و مقالات چاپ شده اخیراً [۷۰ و ۷۱] همچنین از تکنیک‌های هیستوگرام برای شناسایی مجدد شی در محیط‌های با چند دوربین بدون دید مشترک استفاده کرده‌اند. بعضی دیگر از محققان از ویژگی‌های دیگر به همراه هیستوگرام استفاده نموده‌اند. مانند [۶۳] که از ترکیب ویژگی حرکت به همراه هیستوگرام شی استفاده نموده است. [۷۲] از احتمالات برپایه شبکه باور بیزین^۲ (BBN) برای بالابردن عملکرد هیستوگرام برای شناسایی و شناسایی مجدد اشیاء استفاده کرده‌اند. آنها هیستوگرام شی فعلی را با هیستوگرام‌های دیگر ذخیره شده در مجموعه داده انطباق داده‌اند و بعد از محاسبه شباهت بین هیستوگرام‌ها، اگر شباهت چند هیستوگرام بیشتر از حد آستانه باشد در این قسمت نوبت به استفاده از روش بر پایه BBN می‌رسد که شیی که بیشترین احتمال را دارد انتخاب می‌شود.

[۷۳] از بیزین برای ردیابی انسان در چندین دوربین با دید مجزا استفاده کرده است که پاسخ بهینه، یک مجموعه از مسیرهای اشیاء با بیشترین احتمال پسین از مشاهدات بدست آمده، می‌باشد. در این

¹ Signature

² Bayesian Belief Network (BBN)

تحقیق فرض شده که سرعت اشیاء به صورت یکنواخت است و اگر اشیاء، برای چند لحظه متوقف شوند سیستم قادر به شناسایی افراد در دوربین بعدی نیست.

بعضی از محققان از ویژگی بافت برای شناسایی و ردیابی افراد استفاده کرده‌اند [۷۴ و ۷۵]. اما در کل، به دلیل کیفیت پایین دوربین‌های نظارتی فقط استفاده از اطلاعات استخراج شده از بافت برای ردیابی صحیح مناسب نیست [۷۲].

بعضی از کارها، از نقاط کلیدی برای برقراری ارتباط بین اشیاء استفاده کرده‌اند [۷۶] از SIFT برای ردیابی ماشین بکار برده است. [۷۷] نمایشی از SIFT برای شناسایی مجدد انسان استفاده کرده است. [۷۶] نقاط کلیدی را با استفاده از PCA-SIFT [۷۸] در مقیاس بزرگ از دوربین‌ها کاهش داده است. کاهش نقاط کلیدی ارتباط داده‌های بین دوربین‌ها را کاهش می‌دهد ولی عملکرد شناسایی اشیاء نیز کاهش پیدا می‌کند. [۷۹] از SURF اصلاح‌شده و انطباق بین توصیف‌گرها^۱ با استفاده از اولین bin بهتر^۲ و جستجو KD-tree استفاده کرده است.

از دیگر روش‌های شناسایی افراد استفاده از طریقه راه رفتن افراد است. یک روش مخصوص یا وضعیت از حرکت کردن، را راه رفتن^۳ گویند. بعضی از کارهای ابتدایی در راه رفتن و توانایی شناسایی انسان در [۸۰ و ۸۱] بحث شده است. در اوایل، مطالعه به وسیله [۸۱] مشخص کرد که راه رفتن انسان توانایی مفید بیومتریک^۴ برای شناسایی انسان می‌باشد. در کارهای محققان دیگر نیز از راه رفتن برای کاربردهای نظارتی استفاده شده است [۸۲ و ۸۳]. یک روش برای ردیابی و شناسایی انسان بین دوربین‌های متفاوت بدون هم‌پوشانی برپایه آنالیز راه رفتن انسان در [۸۴] معرفی شده است. برای استخراج ویژگی‌های راه رفتن، آنها از مدل‌های حرکت برپایه اطلاعات قسمت‌های بیومتریک اشیاء که توصیف کننده زاویه حرکت از زانو و مفصل ران در حالت‌های متفاوت از دوره‌ای از راه رفتن می‌باشد

^۱ descriptors

^۲ Best Bin First (BBF)

^۳ Gait

^۴ biometric

استفاده کرده‌اند. مقاله [۸۵] نحوه راه رفتن افراد را در هر یک از دوربین‌ها آنالیز نموده و به عنوان معیاری برای برقراری ارتباط بین یک فرد در نظر می‌گیرد. استفاده از راه رفتن نیاز به پردازش نسبتاً سنگینی نیاز دارد و همچنین نمونه‌های ویدئویی در شرایط کاملاً کنترل شده در یک محیط و توسط یک دوربین تهیه شده‌اند. یکی دیگر از خواص بیومتری می‌توان به چهره انسان اشاره کرد که در [۸۶] یک سیستم استریو برای استخراج عمق اشیاء استفاده شده و در کنار آن از رنگ و چهره افراد برای برقراری ارتباط بین مشاهدات مربوط به یک فرد بهره‌گرفته شده است. ویژگی‌های رنگی یک فرد در این کار به سه قسمت تقسیم شده است. رنگ صورت افراد که برای تشخیص چهره استفاده می‌شود، رنگ موی افراد و رنگ سایر قسمت‌های تشخیص داده شده برای فرد توسط سیستم استریو. استفاده از چهره، افراد را ملزم به حرکت در فاصله‌ی کمی از دوربین‌ها می‌نماید تصاویری برای این کار استفاده می‌شوند باید با کیفیت بسیار بالای باشند. در [۸۷] ویژگی چهره به همراه رنگ و بافت لباس افراد برای ردیابی مورد استفاده قرار گرفته است.

[۸۸] یک مدل ظاهری جدید بر پایه کواریانس فضایی (مکانی) ناحیه‌ها که از قسمت‌های مختلف بدن استخراج شده معرفی کرده است. یک شمای هرمی جدید برای برقراری ارتباط قسمت‌های مختلف بدن انسان برای جداکردن ویژگی‌های متمایز انسان معرفی شده است (هر قسمت با بردار شناسایی ۱۱ بعدی مدل شده است).

۴-۲- جمع‌بندی

در این فصل به مرور اجمالی به سیستم‌های ردیابی اشیاء و انسان پرداخته‌ایم. در مورد مشکلات و پیچیدگی‌های این نوع از سیستم‌ها در این فصل بحث شده است. روش‌های موجود در ردیابی انسان را به سه قسمت ردیابی بر پایه مدل حرکت، مدل هندسی و مدل ظاهری معرفی کرده و در ادامه سیستم‌های ردیاب انسان در شبکه از دوربین‌ها با زاویه دید هم‌پوشان و بدون هم‌پوشانی (گسسته) به عنوان دو موضوع اساسی در تحقیقات محققان دهه اخیر معرفی شد و با توجه به نقاط ضعف دوربین‌ها

با دید هم‌پوشان در این فصل و پروژه بیشتر تمرکز به سیستم‌های با زاویه دید مجزا پرداخته شده است. در این فصل به معرفی ویژگی‌های موجود در این نوع از سیستم‌ها و روش‌های موجود در شناسایی مجدد افراد که به عنوان یکی از بخش‌های مهم در این نوع از سیستم‌ها می‌باشد پرداخته شده است. با توجه به نقاط ضعفی که از روش‌های مبتنی بر مدل حرکت و مدل هندسی در شناسایی مجدد افراد در ردیابی انسان در شبکه از دوربین‌ها با دید مجزا مطرح شده، روش‌های مبتنی بر ظاهر را به عنوان روشی قدرتمند در این نوع از سیستم‌ها معرفی کرده‌ایم و در فصل‌های بعدی تمرکز بیشتر بر استخراج ویژگی از مدل ظاهری افراد انجام می‌گیرد.

فصل سوم

حذف پس زمینه و آشکار سازی هدف

۳-۱- مقدمه

آشکارسازی هدف به عنوان یک بخش مهم و تاثیرگذار در سیستم ردیابی می‌باشد که کارایی مراحل بعدی این سیستم به این مرحله وابستگی زیادی دارند. در آشکارسازی، ناحیه‌ی مربوط به هدف مورد نظر در تصویر قطعه‌بندی می‌شود. این مرحله با توجه به شرایط دینامیکی محیط مانند تغییر روشنایی، سایه‌ها، حرکت متفاوت و انسداد بین افراد با چالش‌های زیادی روبرو می‌باشد. در این فصل ابتدا روش‌های موجود بر آشکارسازی هدف که به سه دسته کلی تفاضل فریم‌های متوالی، تفاضل مدل پس‌زمینه و جریان نوری تقسیم شده است و مزایا و معایب هر روش با توجه به کاربرد آن مورد تحلیل قرار گرفته می‌شود. روش تفاضل مدل پس‌زمینه با توجه به اینکه در محیط‌های بسته عملکرد بهتری از لحاظ دقت و زمان محاسبات نسبت به روش‌های موجود دیگر داشته، برای بدست آوردن نواحی متحرک در این پایان‌نامه استفاده شده است. به منظور بالابردن کیفیت هدف استخراج شده از روش‌های حذف سایه و فیلترهای مورفولوژی برای کاهش نویز استفاده شده است. در نهایت پیاده‌سازی این روش‌ها بر روی مجموعه داده معرفی شده در این پایان‌نامه مورد ارزیابی قرار گرفته شده است.

۳-۲- روش‌های آشکار سازی و تشخیص هدف

تشخیص اهداف متحرک یا تشخیص و آشکارسازی حرکت برای تشخیص هدف، اهمیت زیادی در ردیابی بخصوص کاربردهای مراقبتی دارد و حجم زیادی از تلاش‌های تحقیقاتی در زمینه تشخیص و ردیابی در دهه‌های گذشته به این مسئله معطوف گشته است. روش‌های تشخیص اهداف مبتنی بر حرکت بر اساس تغییرات محتویات فریم‌های ویدیویی اهداف را استخراج و آنها را از پس‌زمینه جدا می‌نماید. این روش‌ها بر اساس حرکت به سه دسته تقسیم می‌شود [۸۹]. ۱- تفاضل فریم‌های متوالی^۱
۲- حذف پس‌زمینه^۲ ۳- جریان نوری^۳.

^۱ - Temporal differencing^۲ - Background Subtraction^۳ - Optical Flow

در دو روش نخست فرض بر این است که پس‌زمینه دارای حرکت نمی‌باشد یا عبارت دیگر پس‌زمینه تصویر ثابت^۱ می‌باشد. در این دو روش صلب بودن یا غیرصلب بودن هدف اهمیتی ندارد و در هر دو صورت هدف آشکار خواهد شد، اما روش جریان نوری در مواردی که پس‌زمینه ثابت نباشد یا دارای حرکت کلی نیز باشد، نیز قابل استفاده است. ولی اشیاء در این روش باید صلب باشند. و اجزاء مختلف شیء دارای مولفه‌های حرکتی ثابتی نسبت به مدل حرکتی پس‌زمینه باشند. علاوه بر این، روش جریان نوری نسبت به دو روش قبل از بار محاسباتی بسیار بالاتری برخوردار است و اغلب از این روش بعثت بار محاسباتی سنگین آن استفاده نمی‌شود. در ذیل این سه روش را با تفصیل بیشتری بیان خواهیم کرد.

۳-۲-۱ روش تفاضل فریم‌های متوالی

روش تفاضل فریم‌های متوالی نسبت به تغییرات نوری محیط تطبیق‌پذیر بوده و عملکرد خوبی را از خود نشان می‌دهد. این بدین معنی است، اگر تغییرات زائدی مابین دو فریم متوالی وجود داشته که جز شی متحرک (هدف) نباشد، این روش در برابر این تغییرات مقاوم بوده و تغییرات واقعی که از هدف مورد نظر ایجاد شده را آشکار می‌نماید، ولی در عوض در اکثر موارد در استخراج همه نقاط مربوط به هدف ضعیف عمل می‌کند و معمولاً فقط نقاط مربوط به لبه‌های هدف را آشکار می‌سازد [۹۰].

در این روش از تکنیک اعمال آستانه مناسب به تصاویر حاصل از تفاضل فریم‌های متوالی استفاده می‌شود. خروجی این روش فریم‌ها را به دو ناحیه ثابت (اغلب پس‌زمینه) و متحرک (اغلب پیش‌زمینه) تفکیک می‌کند. برای توضیح این روش، فرض کنید که $I_n(x)$ نشان دهنده شدت روشنایی فریم در نقطه x در فریم n ام باشد. آنگاه فریم تفاضل بصورت رابطه (۳-۱) تعریف می‌شود:

$$FD(x) = I_n(x) - I_{n-1}(x) \quad (۳-۱)$$

^۱ - Fixed

فریم تفاضل در حقیقت تفاضل نقطه به نقطه بین دو فریم است. با فرض اینکه شرایط نوری محیط تقریباً بین دو فریم متوالی ثابت می‌باشد، مقدار فریم تفاضل در یک نقطه بیانگر تغییر در یک نقطه می‌باشد. مقدار غیرصفر تفاضل به دو دلیل ایجاد می‌شود. نخست به دلیل حرکت اشیاء موجود در تصویر، و دوم به دلیل نویز مشاهده^۱ بین مقادیر یک نقطه در دو فریم متوالی.

به‌منظور تمایز بین تغییرات ایجادشده بوسیله نویز و تغییرات بر اثر حرکت اشیاء، از روش جداسازی^۲ مبتنی بر آستانه^۳ استفاده می‌کنیم. در این حالت فرض می‌شود که نویز مشاهده یک نویز گوسی با متوسط صفر می‌باشد. در نتیجه این فرض می‌توان هر نقطه از تصویر را بصورت یک متغیر تصادفی گوسی با متوسط غیرصفر و واریانس σ در نظر گرفت.

در نتیجه مقدار هر نقطه در فریم تفاضل بصورت یک متغیر تصادفی با متوسط صفر و واریانس $\sqrt{2}\sigma$ خواهد بود. قدر مطلق مقدار هر نقطه در فریم تفاضل با یک آستانه مقایسه می‌شود. چنانچه این مقدار از این آستانه بیشتر باشد، آن نقطه بعنوان نقطه متحرک یا پیش‌زمینه تلقی خواهد شد و چنانچه این مقدار از آستانه کمتر باشد، این نقطه به‌عنوان پس‌زمینه تلقی خواهد شد. این پردازش بصورت زیر نمایش داده می‌شود:

$$Binary_Mask_n(\mathbf{x}) = \begin{cases} 1 & \text{if } |FD_n(\mathbf{x})| > T_n(\mathbf{x}) \\ 0 & \text{other wise} \end{cases} \quad (2-3)$$

در رابطه (۲-۳)، $T_n(\mathbf{x})$ ، آستانه تصمیم‌گیری برای فریم n ام در نقطه $\mathbf{x}$ می‌باشد. این آستانه از توزیع آماری تغییرات مقدار روشنایی یک نقطه و میزان تشخیص اشتباه نقاط بدست می‌آید.

هنگام کار ممکن است که نقاط منفرد^۴ به‌علت وجود نویز در خروجی پردازش ایجاد شود. این نقاط را می‌توان بوسیله پس‌پردازش^۵ حذف کرد. بعنوان مثال با محاسبه پیوستگی^۴ تایی^۱ یا ۸ تایی^۲ نقاط

^۱ - Observation noise

^۲ - Segmentation

^۳ - Threshold

^۴ - Isolated

^۵ - Post-Processing

باینری، پیش‌زمینه را به گروه‌های متصل (حباب) تفکیک می‌کنیم و اگر تعداد نقاط در یک حباب (مساحت حباب) از تعداد مشخصی کمتر باشد، این گروه به عنوان نویز تلقی شده و حذف می‌شود. در روش دیگر، به منظور حذف این نقاط منفرد از روش تفاضل تجمعی^۳ که از حافظه بیشتری در تشخیص حرکت استفاده می‌کند، استفاده می‌شود. برای این منظور فرض کنید، که $I_n(x)$ فریم جاری و $N, I_{n-N}(x), \dots, I_{n-2}(x), I_{n-1}(x)$ فریم ماقبل آن باشد. آنگاه با اعمال تفاضل بین فریم جاری و هر کدام از فریم‌های قبلی، N فریم تفاضل بدست می‌آوریم. پس از اعمال سطح آستانه به این N فریم تفاضل و بدست آوردن N فریم باینری، اگر تعداد نقاط روشن برای مختصات یک نقطه در N فریم باینری متوالی از حدی بیشتر شود، آنگاه آن نقطه بعنوان نقطه پیش‌زمینه در فریم باینری حاصل، لحاظ خواهد شد. در حقیقت هر چند بار که یک نقطه برای تعداد بار بیشتری در N فریم باینری به-عنوان پیش‌زمینه تلقی شود، احتمال تعلق آن نقطه به پیش‌زمینه بیشتر خواهد بود و احتمال وجود آن بر اثر نویز کمتر خواهد شد. ولی در عوض، تاخیر و حافظه بیشتر از مشکلات آن بشمار می‌آید. برای مثال برای دو فریم متوالی، حاصل بدین‌گونه محاسبه خواهد شد:

(۳-۳)

$$Binary - Mask_n(x) = \begin{cases} 1 & \text{if } (|I_n(x) - I_{n-1}(x)| > T_n(n)) \text{ and } (|I_n(x) - I_{n-2}(x)| > T_n(x)) \\ 0 & \text{otherwise} \end{cases}$$

۳-۲-۲- روش حذف پس زمینه

در بسیاری از کاربردهای مراقبت ویدئویی، دوربین‌ها ثابت هستند و ما به تشخیص اشیاء پیش‌زمینه علاقه‌مند هستیم. این اشیاء اغلب شامل اتومبیل‌ها یا افراد می‌باشد. مشخصاً، روش تفاضل فریمی، آگاهی ناقصی در مورد نواحی که ممکن است حاوی اشیاء متحرک باشد به ما می‌دهد و همانگونه که ذکر شد، نمی‌تواند نواحی پس‌زمینه و پیش‌زمینه را به‌گونه به‌یکباره از هم جدا کند و در نتیجه

^۱- 4 Connected Component

^۲- 8 Connected Component

^۳- Accumulative Difference

نمی‌تواند در بسیاری از موارد نتایج مورد دلخواه فرآیند ردیابی را فراهم آورد.

در روش حذف پس‌زمینه، برای جدا کردن ناحیه پیش‌زمینه از ناحیه پس‌زمینه از یک مدل آماری که توصیف‌کننده احتمالی مقادیر پس‌زمینه است، استفاده می‌شود و در حقیقت هر نقطه از تصویر به یکی از دو مجموعه پس‌زمینه یا پیش‌زمینه دسته‌بندی خواهد شد. در اینجا فرض می‌شود که دوربین ثابت است و ما می‌توانیم یک مدل آماری یا توزیع احتمالی برای هر نقطه از پس‌زمینه بدست آوریم. اغلب چنین مدلی از پس‌زمینه، نقطه‌ای می‌باشد. بدین معنی که هر نقطه به‌گونه مستقل و مجزا مدل می‌شود. یا به عبارت دیگر هر نقطه دارای تابع توزیع احتمال مخصوص به خود خواهد بود. برای هر نقطه از تصویر، مقدار روشنایی یا رنگ آن نقطه از تصویر با مدل پس‌زمینه مقایسه می‌شود و بر این اساس این مقایسه نوع نقطه اعم از پیش‌زمینه یا پس‌زمینه بودن آن، مشخص می‌گردد. بعنوان مثال، در یک روش ساده می‌توان یک تصویر میانگین را بعنوان مدل پس‌زمینه گزینش نمود و نقاطی که فاصله مقادیر روشنایی آنها از مقادیر متناظر در مدل بیشتر از آستانه بخصوصی باشد، بعنوان پیش‌زمینه گزینش خواهند شد. مدل پس‌زمینه را می‌توان از یک دنباله از فریم‌ها بدست آورد. مدل‌های آماری مختلفی برای پس‌زمینه استفاده می‌شود که از مهمترین آنها می‌توان به موارد زیر اشاره کرد.

• مدل میانگین^۱

برای هر نقطه (x, y) ، یک مقدار میانگین $B(x, y)$ بعنوان مقدار پس‌زمینه استفاده می‌شود. برای تعیین نوع نقطه، از لحاظ تعلق به پس‌زمینه یا پیش‌زمینه از شرط زیر استفاده می‌کنیم. اگر رابطه ذیل (۳-۴) برقرار باشد، نقطه جزء پیش‌زمینه است وگرنه متعلق به پس‌زمینه خواهد بود.

$$|I(x, y) - B(x, y)| > T \quad (۳-۴)$$

در رابطه فوق T مقدار آستانه تصمیم‌گیری و $I(x, y)$ مقدار روشنایی تصویر ورودی در نقطه مزبور می‌باشد. مدل پس‌زمینه باید در طول زمان برای تطبیق با تغییرات تدریجی روشنایی محیط و سایر

^۱- Mean

عوامل، خود را وفق دهد. یکی از مدل‌هایی که برای تطبیق مدل پس‌زمینه در این حالت ارائه شده بصورت زیر است.

$$B_{n+1}(x) = \begin{cases} \alpha B_n(x) + (1-\alpha)I_n(x) & \text{X is non - moving} \\ B_n(x) & \text{X is moving} \end{cases} \quad (5-3)$$

در این مدل سطح آستانه نیز مورد تطبیق قرار می‌گیرد:

$$T_{n+1}(x) = \begin{cases} \alpha T_n(x) + (1-\alpha)(5 \times |I_n(x)| - B_n(x)) & \text{X is non - moving} \\ T_n(x) & \text{X is moving} \end{cases} \quad (6-3)$$

در روابط (5-3) و (6-3)، α نرخ یا سرعت تطبیق مدل با تغییرات محیط می‌باشد. اگر $\alpha = 0$ ، اختیارشود، بیشترین تطبیق با تغییرات اعمال می‌شود و هر قدر که α به یک نزدیکتر باشد، سرعت تطبیق‌پذیری کمتر خواهد بود. از α بعنوان حافظه سیستم نیز یاد می‌شود. روابط فوق بصورت فیلتر خطی با پاسخ ضربه نامحدود^۱ می‌باشد. $B_n(x)$ در رابطه (5-3) بصورت متوسط خطی مقادیر روشنایی نقاط تصویر $I_n(x)$ است. در حالیکه، $T_n(x)$ در (6-3) به سمت پنج برابر انحراف معیار مقادیر $I_n(x)$ میل می‌کند.

• مدل میانگین و کواریانس^۲

برای توصیف هر نقطه از تصویر از توزیع گوسی استفاده می‌شود [۹۱-۹۳]. یا به عبارت دیگر توزیع آماری هر نقطه بوسیله یک میانگین $B(x,y)$ و کواریانس آن، $C(x,y)$ توصیف خواهد شد. برای بررسی تعلق یک نقطه $I(x,y)$ ، به پس‌زمینه از فاصله ماحالانوبیس^۳ بصورت رابطه (۷-۳) استفاده خواهد شد. نقطه متعلق به پیش‌زمینه رابطه زیر را ارضاء خواهد کرد.

$$[I(x,y) - B(x,y)]^T C^{-1}(x,y) [I(x,y) - B(x,y)] > T \quad (7-3)$$

برای بروز رسانی پارامترها برای مدل پس‌زمینه از رابطه‌های (۸-۳) و (۹-۳) استفاده می‌شود.

^۱ - Infinite Impulse Response (IIR)

^۲ - Mean & Covariance

^۳ - Mahalanobis

$$\mu_{t+1} = \alpha\mu_t + (1-\alpha)Z_{t+1} \quad (۸-۳)$$

$$\sigma_{t+1}^2 = \alpha(\sigma_t^2 + (\mu_{t+1} - \mu_t)^2) + (1-\alpha)(Z_{t+1} - \mu_{t+1})^2 \quad (۹-۳)$$

که Z_{t+1} مقدار پارامترهای در نظر گرفته شده برای پس‌زمینه در زمان $t+1$ است و مقدار ثابت α نرخ بروز رسانی را با توجه به تغییرات صحنه کنترل می‌کند. اگر این مقدار ثابت کوچک انتخاب شود، داده‌های قدیمی سریع فراموش می‌شوند.

• انحراف زمانی^۱

در این مدل برای هر نقطه مقدار ماکزیمم شدت روشنایی آن نقطه، $M(x, y)$ ، و مینییمم مقدار شدت روشنایی آن نقطه، $N(x, y)$ ، و $D(x, y)$ بزرگترین قدرمطلق تفاضل مقادیر روشنایی در آن نقطه بین زوج فریم‌های متوالی نگهداری می‌شود [۹۴]. اگر مقدار روشنایی در این نقطه، $I(x, y)$ در تصویر ورودی شروط زیر را ارضا کند، بعنوان پیش‌زمینه تلقی می‌شود.

$$|I(x, y) - N(x, y)| > D(x, y) \quad (۱۰-۳)$$

$$|I(x, y) - M(x, y)| > D(x, y) \quad (۱۱-۳)$$

• مدل گاوسی مخلوط^۲

در این روش از چند تابع گاوسی برای تخمین توزیع مقدار هر پیکسل استفاده می‌شود [۹۵ و ۹۶]. برای این کار m تابع گاوسی برای محاسبه‌ی میزان تعلق یک پیکسل، x به یک شیء O (در این جا پس زمینه) بکار گرفته شده است.

$$P(x|O) = \sum_{j=1}^m P(x|j)\pi(j) \quad (۱۲-۳)$$

که $\pi(j)$ وزن هر تابع گاوسی است و $\sum_{j=1}^m \pi(j) = 1$. هر تابع گاوسی $P(x|j)$ نیز به صورت زیر

تعریف می‌شود:

^۱ Temporal Deviation

^۲ Mixture of Gaussian Model

$$P(x | j) = \frac{1}{2\pi \left| \sum_j \right|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_j)^T \sum_j^{-1} (x-\mu_j)} \quad (۱۳-۳)$$

که μ و $\sum$ مقدار میانگین و ماتریس کواریانس تابع گاوسی هستند. برای بدست آوردن پارامترهای هر تابع گاوسی مخلوط اولیه برای هر پیکسل از روی مجموعه آموزشی $X^{(0)}$ با اندازه $N^{(0)}$ از الگوریتم حداکثر احتمال وقوع^۱ (EM) استفاده می‌شود. الگوریتم (EM) به صورت تکراری پارامترهای مدل را بروز رسانی می‌کند. اگر $P^{old}(j | X)$ مجموع احتمال پسین^۲ فرض شود پارامترهای مدل به صورت زیر در هر مرحله بروز رسانی می‌گردد:

$$\mu_j^{new} = \frac{\sum_{x \in X^{(0)}} P^{old}(j | X) X}{\psi^{old}} \quad (۱۴-۳)$$

$$\pi_j^{new} = \frac{\psi_j^{old}}{\pi_j^{old}} \quad (۱۵-۳)$$

$$\sum_j^{new} = \frac{\sum_{x \in X^{(0)}} P^{old}(j | X) (x - \mu_j^{new})^T (x - \mu_j^{new})}{\psi^{old}} \quad (۱۶-۳)$$

$$P(j | X) = \frac{P(X | j) \pi(j)}{P(X | O)} \quad (۱۷-۳)$$

الگوریتم EM به صورت یکنواخت مقدار احتمال را در هر تکرار افزایش می‌دهد تا به بیشینه محلی همگرا شود. نتیجه حاصل از مدل گاوسی مخلوط به تعداد توابع گاوسی m و مقدار دهی اولیه پارامترهای تابع گاوسی وابسته است. میانگین توابع گاوسی با انتخاب یکی از نقاط مجموعه‌ی آموزشی به صورت تصادفی و وزن هر تابع گاوسی π نیز با مقدار $\frac{1}{m}$ مقدار دهی اولیه شده است.

• مدل گاوسی مخلوط تطبیقی^۳

^۱ Expectation Maximization

^۲ Posterior Probability

^۳ Adaptive Mixture of Gaussian Model

در این روش از چند تابع گاوسی برای تخمین توزیع مقدار هر پیکسل استفاده می‌شود. احتمال اینکه یک پیکسل مطمئناً مقدار X_n در زمان N داشته باشیم می‌توان نوشت:

$$P(X_N) = \sum_{j=1}^N w_j \eta(X_N; \theta_j) \quad (۱۸-۳)$$

که w_j وزن هر تابع گاوسی است. هر تابع گاوسی $\eta(X; \theta_k)$ نیز به صورت زیر تعریف می‌شود:

$$\eta(X, \theta_k) = \eta(X; \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(X-\mu_k)^T \Sigma_k^{-1} (X-\mu_k)} \quad (۱۹-۳)$$

که μ_k و $\Sigma_k = \sigma_k^2 I$ مقدار میانگین و ماتریس کواریانس تابع گاوسی هستند. K توزیع برحسب مقدار $\frac{w_k}{\sigma_k}$. توزیع اول به عنوان مدل زمینه مورد استفاده قرار می‌گیرد که مدل زمینه به صورت رابطه (۲۰-۳) تخمین زده می‌شود:

$$B = \arg \min_b \left(\sum_{k=1}^b w_k > T \right) \quad (۲۰-۳)$$

T در رابطه (۲۰-۳) حدآستانه می‌باشد. در واقع کمترین درصد از مدل پس‌زمینه می‌باشد و در واقع وزن k را تا جایی اضافه می‌نماییم که به مقدار حدآستانه برسد. توزیع B برای تمامی پیکسل‌ها محاسبه می‌شود و پیکسلی که مقدار آن با مقدار B منطبق بود به عنوان زمینه در نظر گرفته می‌شود. بعد از بدست آوردن مدل پس‌زمینه باید پارامترهای تابع گاوسی در این مدل به روز رسانی شوند. μ و Σ پارامترهای هر تابع گاوسی می‌باشد که باید در هر مرحله بروز رسانی شوند. در صورتی که مقدار پیکسل X_N با k توزیع مربوط به آن مقایسه و مقدار پیکسل در رابطه (۲۱-۳) صدق کند، پارامترها مدل باید به روز شود.

$$\frac{(X_N - \mu_N)}{\sigma_N} > 2.5 \quad (۲۱-۳)$$

برای بروز رسانی پارامترها در صورتی که در شرط رابطه (۲۱-۳) صدق کند از رابطه‌های زیر استفاده می‌نماییم.

$$\hat{w}_k^{N+1} = (1-\alpha) \hat{w}_k^N + \alpha \hat{P}(\omega_k | X_{N+1}) \quad (22-3)$$

$$\hat{\mu}_k^{N+1} = (1-\alpha) \hat{\mu}_k^N + \alpha \rho X_{N+1} \quad (23-3)$$

$$\hat{\Sigma}_k^{N+1} = (1-\alpha) \hat{\Sigma}_k^N + \rho (X_{N+1} - \hat{\mu}_k^{N+1})(X_{N+1} - \hat{\mu}_k^{N+1})^T \quad (24-3)$$

در رابطه‌های بالا، α ضریب یادگیری و $\hat{P}$ در صورت تطبیق مقدار پیکسل با انحراف معیار، مقدار ۱ و در غیر این صورت مقدار صفر بر می‌گرداند.

$$\rho = \alpha \eta(X_{N+1}; \hat{\mu}_k^N, \hat{\Sigma}_k^N) \quad (25-3)$$

$$\hat{P}(\omega_k | X_{N+1}) = \begin{cases} 1 & \text{; if } \omega_k \text{ is the first match Gaussian component} \\ 0 & \text{; otherwise} \end{cases} \quad (26-3)$$

• پیش‌بینی خطی^۱

مدل پیش‌بینی خطی از فیلتر وینر^۲ برای مدل‌کردن الگوی زمانی هر نقطه استفاده می‌کند. در این مدل از p فریم گذشته برای پیش‌بینی فریم جاری استفاده می‌شود. این مدل قادر است، تا خود را با تغییرات روشنایی وفق دهد. پس از تخمین پارامترها برای فیلتر وینر، انجام پردازش تفاضل پس‌زمینه راحت خواهد بود. مدل پیش‌بینی خطی بصورت زیر می‌باشد:

$$B_t(x, y) = -\sum_{k=1}^p a_k B_{t-k}(x, y) \quad (27-3)$$

در محیط بیرون^۳ با پس‌زمینه متغیر دو منبع تشخیص اشتباه^۴ وجود دارد. نخست تشخیص اشتباه برای نویز تصادفی، که در سراسر صفحه ممکن است و دوم تشخیص‌های اشتباه برای حرکت‌های کوچک در پس‌زمینه، که بوسیله مدل پس‌زمینه نمایش داده نمی‌شود. به‌عنوان مثال اگر شاخه درخت

¹ - Linear Prediction

² - Wiener

³ - Outdoor Environment

⁴ - False Detection

بیشتر از آن چیزی که در مدت تشکیل مدل حرکت می‌کرد، تکان بخورد یا جابجایی‌های کوچک در محل دوربین به سبب باد به وجود آید، احتمال وقوع تشخیص‌های اشتباه زیاد خواهد بود.

مرحله دوم در تشخیص اشیاء متحرک حذف تشخیص‌های اشتباه برای حرکت‌های کوچکی است که در مدل در نظر گرفته نشده‌اند. اگر قسمتی از پس‌زمینه (یک شاخه درخت) حرکت کند و یک نقطه جدید را اشغال کند، که مقدار روشنایی جدید این نقطه، در مدل آماری پس‌زمینه نقطه وجود نداشته باشد، این نقطه به عنوان پیش‌زمینه تلقی خواهد شد. اما این نقطه به احتمال قوی متعلق به توزیع پس‌زمینه در مکان اصلی خود خواهد بود. با فرض اینکه جابجایی کوچکی ما بین دو فریم متوالی اتفاق افتاده، می‌توان با در نظر گرفتن توزیع پس‌زمینه در همسایگی نقطه مزبور، تعیین کرد که آیا این نقطه متعلق به پس‌زمینه همسایه بوده و در اثر حرکت پس‌زمینه بوجود آمده، یا واقعا نقطه ای از پیش‌زمینه است.

با اعمال آستانه در نقطه آشکار شده به عنوان پیش‌زمینه، می‌توان بسیاری از تشخیص‌های اشتباه به علت حرکت‌های کوچک در پس‌زمینه را حذف کرد. متأسفانه با انجام این پردازش بعضی از نقاطی که به درستی نیز به عنوان پیش‌زمینه گزینش شده بودند، نیز حذف می‌گردند. زیرا که ممکن است بعضی از نقاط آشکار شده واقعی به عنوان پیش‌زمینه در مرحله قبل، بر حسب تصادف مشابه پس‌زمینه نقاط همسایه باشند. که این مورد اغلب در مورد تصاویری که فقط با سطح روشنایی^۱ نمایش داده می‌شوند، اتفاق می‌افتد. برای پرهیز از دست دادن چنین تشخیص‌های درستی، از شرط اینکه همه نقاط پیش‌زمینه تشخیص داده شده، باید از یک همسایگی مجاور آمده باشند و نه فقط بعضی از نقاط محدود، استفاده می‌کنیم. بدین منظور احتمال اینکه یک مؤلفه متصله^۲، از یک محل همسایگی حرکت کرده باشد، را در نظر می‌گیریم.

۳-۲-۳- روش جریان نوری

^۱- Gray level

^۲ - Connected Component

در این روش یک میدان برداری دو بعدی که تصویر^۱ بردارهای حرکت واقعی اشیاء در سه بعد می‌باشد، استخراج می‌شود. به این میدان برداری، جریان نوری گفته می‌شود. در این روش فرض می‌شود، که اشیاء متحرک در صحنه دارای حرکت‌های مستقلی از حرکت پس‌زمینه هستند. حرکت پس‌زمینه می‌تواند، در اثر حرکت دوربین ایجاد شده باشد. حال با فرض حرکت کلی پس‌زمینه^۲، یک مدل پارامتری برای آن استخراج می‌شود. با استفاده از این مدل پارامتری، بردار حرکت معادل بر حسب حرکت کلی پس‌زمینه در آن نقطه تخمین زده می‌شود. حال با مقایسه این بردار با بردار جریان نوری در آن نقطه، حرکت باقی‌مانده^۳ در آن نقطه محاسبه می‌شود. در صورتیکه بردار حرکت جریان نوری با بردار بدست آمده از حرکت کلی پس‌زمینه همخوانی داشته باشد، بدین معنی خواهد بود، که آن نقطه یا بلوک جزء پس‌زمینه می‌باشد و حرکت باقی‌مانده تقریباً صفر خواهد بود. اما هر چه این همخوانی کمتر باشد، یا تفاوت بین دو بردار بیشتر باشد، نشان‌دهنده وجود شیء متحرک مستقل از پس‌زمینه در آن نقطه یا بلوک خواهد بود. اگر نقاط یا بلوک‌هایی که دارای حرکت متفاوت با حرکت کلی پس‌زمینه هستند را دسته‌بندی کنیم، تصویر به دو دسته نقاط پس‌زمینه و نقاط پیش‌زمینه، تفکیک خواهد شد. حال اگر فرض شود که همه بردارهای حرکت مربوط به یک شیء در یک سو و راستا و تقریباً هم‌اندازه باشند یا به تعبیر دیگر شیء دارای حرکت صلب باشد، می‌توان با دسته‌بندی این بردارهای حرکت که دارای مشابهت هستند، به همه ناحیه شیء دسترسی پیدا کرد یا در حقیقت می‌توان گفت که شیء متحرک را از بقیه جدا کرده‌ایم [۹۷].

اما در صورتیکه شیء دارای حرکت صلب نباشد یا به عبارت دیگر اجزاء مختلف یک شیء دارای بردارهای مختلف حرکتی باشند، این دسته‌بندی براساس شباهت بردارها، منجر به جداسازی و تشخیص شیء نخواهد شد و فقط از معیار انحراف از حرکت پس‌زمینه برای تشخیص نواحی متحرک

¹- Projection

²-Global Motion

³- Residual Motion

می‌توان استفاده نمود.

با توجه به مطالب فوق، استفاده از یک روش مناسب برای حذف پس‌زمینه و آشکارسازی هدف می‌تواند تاثیر بسیار زیاد در عملکرد سیستم ردیابی انسان داشته باشد به همین منظور جدول (۳-۱) مقایسه‌ای از روش‌های مطرح شده برای حذف پس‌زمینه را نشان می‌دهد و نقاط قوت و ضعف هر روش را با توجه به مطالب این فصل، بیان می‌کند.

جدول (۳-۱). مقایسه روش‌های آشکارسازی

الگوریتم	فوائد	مشکلات
اختلاف فریم	نرخ تشخیص بهتر و سرعت آشکارسازی بالاتر	تشخیص اشتباه شی در پس زمینه های پیچیده
جریان نوری	پایداری	بار محاسباتی زیاد و نیاز به پردازش بلادرنگ
تفریق پس زمینه	کارایی زیاد و دقت تشخیص بالا	پیچیدگی

۳-۳- حذف سایه

تشخیص سایه اشیاء به‌عنوان نواحی پیش‌زمینه یکی از دلایل اشتباه و سردرگمی در مراحل بعدی پردازش می‌باشد. ما علاقه‌مند هستیم که در خروجی پردازش حذف پس‌زمینه بین اهداف و سایه‌های آنها تمایز قائل شویم. داده‌ها مربوط به رنگ نقاط در حذف سایه بسیار مفید است. زیرا در نقاطی که سایه اشیاء یا اهداف قرار دارد، این نقاط دارای مقدار روشنایی کمتری نسبت به نقاط دیگر پس‌زمینه یا همسایه می‌باشند. ولی رنگ این نقاط تقریباً ثابت است و با استفاده از یک مختصات نمایش رنگ می‌توان سایه را تاحدی حذف کرد.

در الگوریتمی که ذکر می‌شود، با جدا کردن داده‌های رنگ از داده‌های روشنایی می‌توان سایه را حذف

نمود. برحسب سه متغیر رنگ R,G,B داده شده مختصات رنگی r^1 , g, b بصورت زیر تعریف می‌شود:

$$r = \frac{R}{R+G+B} \quad (28-3)$$

$$g = \frac{G}{R+G+B} \quad (29-3)$$

$$b = \frac{B}{R+G+B} \quad (30-3)$$

که این مختصات جدید رابطه $r+g+b=1$ را ارضاء می‌کند. یا به عبارت دیگر به این مختصات نرمالیزه نیز می‌گوییم. مزیت استفاده از مختصات رنگی فوق در این است که نسبت به تغییرات کوچک در شرایط نوری محیط بخاطر وجود سایه غیرحساس است. توجه داشته باشید که حذف پس‌زمینه با روش ذکر شده در هر مختصات رنگی امکان‌پذیر است. برای حذف سایه از مختصات رنگ (r, g) استفاده می‌کنیم. در این پایان نامه از این روش برای حذف سایه استفاده شده است.

روش دیگر استفاده از مختصات رنگ (HSI) می‌باشد. در این مختصات نیز ابتدا پردازش حذف پس‌زمینه روی مقادیر I انجام می‌پذیرد و سپس شرط وجود سایه با مقادیر S,H بررسی می‌گردد. در این مختصات رنگ نیز در صورت وجود سایه مقدار روشنایی نقطه، تغییر می‌کند ولی مقدار رنگ در آن نقطه، تقریباً ثابت می‌ماند.

مشابه آنچه که در الگوریتم تفاضل فریم‌های متوالی ذکر شد، برای حذف نویز و نقاط منفرد، از اعمال فیلتر میانه^۲ و همچنین دسته‌بندی بصورت گروه‌های متصل ۸تایی و ۴تایی و سپس حذف مجموعه‌هایی که تعداد آنها یا به عبارت دیگر مساحت آنها از حدی کمتر باشد، نیز استفاده می‌شود. در روش دیگر حذف نویز از فیلترهای ریخت‌شناسی^۳ مانند سایش^۴ یا بازکردن^۵ نیز می‌توان استفاده کرد.

¹ - Chromaticity Coordinates

² -Median Filter

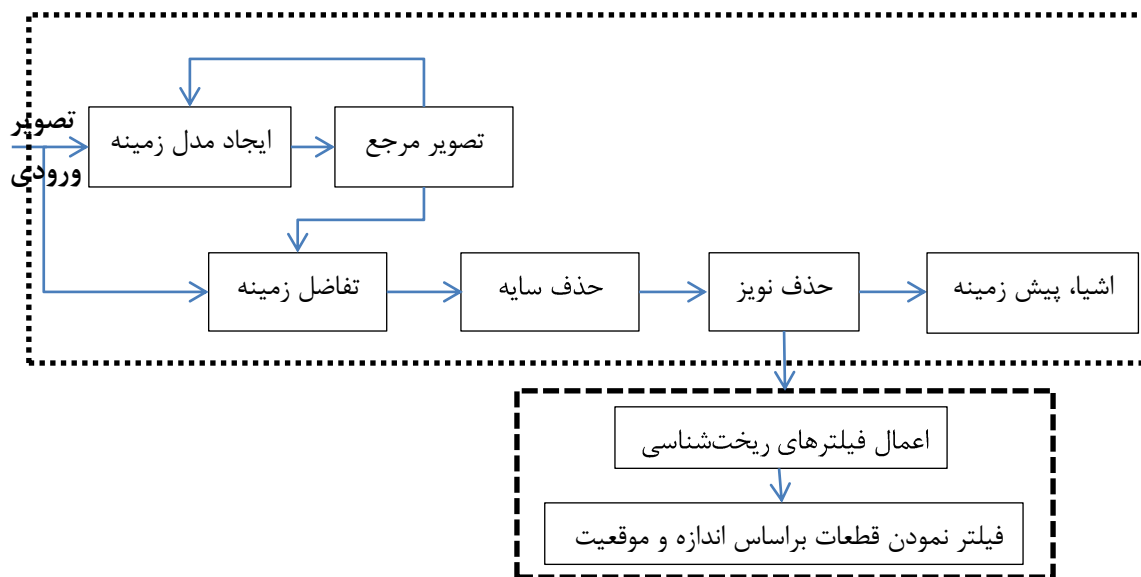
³ -Morphological filters

⁴ -Erosion

⁵ -Opening

۳-۴- روش مورد استفاده در پایان‌نامه به منظور آشکارسازی هدف

همانطور که در بخش پیشین معرفی شد، به دلیل پایین بودن پیچیدگی محاسباتی و مدل کردن پس‌زمینه به نحو نسبتاً خوب، از توزیع گاوسی برای بردار مقادیر هر پیکسل استفاده شده است. به طور کلی، مدل کردن پس‌زمینه برای محیط‌های بسته خوب عمل می‌کند، ولی برای محیط‌های باز اگر تغییرات و حرکات خیلی سریع باشد مشکل ایجاد می‌شود. البته با انتخاب مقادیر مناسب برای هر پیکسل جهت ایجاد مدل گاوسی این مشکل را می‌توان تا حد زیادی کاهش داد. همچنین با انتخاب پارامترهایی مثل گرادیان محلی می‌توان تاثیر سایه را نیز حذف کرد. در این پروژه با توجه به اینکه محیط بسته می‌باشد از مدل گاوسی مخلوط تطبیقی برای تخمین پس‌زمینه استفاده شده است. برای تقطیع حرکت در هر فریم چند مدل گاوسی ایجاد شده بوسیله‌ی فریم‌های قبلی بروز رسانی می‌گردد، تفاضل پس‌زمینه انجام می‌شود. بعد از حذف پس‌زمینه برای از بین بردن سایه از روش مطرح شده به عنوان نرمالیزه (در قسمت حذف سایه) و همچنین از روش‌های ترکیبی ریخت‌شناسی برای از حذف نویز استفاده شده است. از فیلتر فرسایش برای از بین بردن نویزهای ضربه‌ای و از فیلتر بازکردن برای اتصال قسمت‌های مجزای نزدیک به هم، که به دلیل مشکلات قطعه‌بندی به اشتباه از هم جدا شده‌اند استفاده شده است. و در انتها با در نظر گرفتن یک حد‌آستانه براساس اندازه و موقعیت، اهداف از تصاویر ویدیویی استخراج شده اند. شمای کلی برای آشکارسازی افراد در این پروژه به صورت شکل (۳-۱) می‌باشد.



(۱-۳) شمای کلی برای آشکارسازی اهداف در تصاویر ویدیویی

شکل (۳-۳) نتایج بدست آمده از جداسازی پس‌زمینه که تصویر پس‌زمینه در شکل (۲-۳) نشان داده شده بر روی مجموعه داده مورد استفاده را نشان می‌دهد.



شکل (۲-۳) تصویر زمینه برای دو دوربین با زاویه دید مجزا



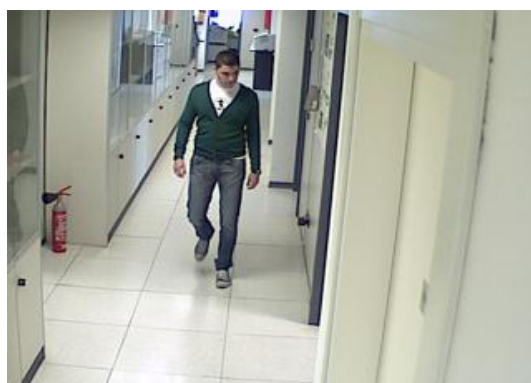
الف-۱



الف-۲



الف-۳



الف-۴



ب-۱



ب-۲

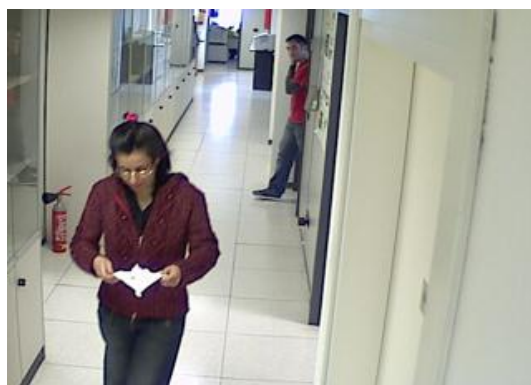


ب-۳



ب-۴





۱-پ



۲-پ



۳-پ



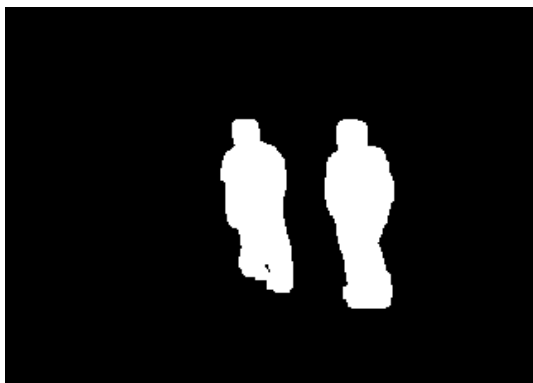
۴-پ



۱-ج



۲-ج



ج-۴



ج-۳



ی-۲



ی-۱



ی-۴



ی-۳

شکل (۳-۳) استخراج افراد از مجموعه مورد استفاده در این پایان‌نامه با استفاده از روش مدل گاوسی ترکیبی مخلوط تطبیقی. (الف-۱)، (ب-۱)، (پ-۱)، (ج-۱) و (ی-۱) یک فریم از توالی حرکتی داده‌های مورد استفاده. (الف-۲)، (ب-۲)، (پ-۲)، (ج-۲) و (ی-۲) تصویر دودویی باینری افراد بعد از تفریق فریم فعلی از مدل پس زمینه. (الف-۳)، (ب-۳)، (پ-۳)، (ج-۳) و (ی-۳) تصویر دودویی باینری افراد بعد از حذف سایه. (الف-۴)، (ب-۴)، (پ-۴)، (ج-۴) و (ی-۴) تصویر دودویی باینری افراد بعد از حذف سایه و حذف نویز.

۳-۵- جمع‌بندی

در این فصل به معرفی روش‌های موجود در شناسایی و آشکارسازی هدف در یک تصویر ویدئویی پرداخت شد و در مورد نقاط ضعف و قوت این روش‌ها بحث شده است. اکثر سیستم‌های نظارت ویدئویی با تشخیص و تقطیع حرکت شروع می‌شوند. هدف این مرحله، تقطیع نواحی مربوط به اشیاء متحرک از بقیه تصویر می‌باشد. مراحل بعدی کار بسیار وابسته به مرحله‌ی تقطیع حرکت می‌باشد. در این مرحله باید صحنه یا محیط به خوبی مدل شود تا حساسیت نسبت به تغییراتی مثل تغییرات نور، سایه‌ها حداقل کاهش یابد. پس از ایجاد مدل برای محیط، باید در طول کار نظارت ویدئویی، مدل بروز رسانی گردد. سپس با تفریق تصویر صحنه از مدل پس زمینه، اشیاء متحرک موجود در صحنه قابل استخراج هستند. به طور کلی قسمت‌های دیگر سیستم نظارت ویدئویی به قسمت تقطیع حرکت وابسته است، به طوری که اگر عمل تقطیع حرکت خوب انجام نگیرد، کارایی قسمت‌های دیگر بسیار پایین می‌آید. در این فصل ابتدا روش‌های مختلف جداسازی پس‌زمینه را بررسی کردیم و سپس روش مورد استفاده‌ی این پروژه را مرور کرده و نتایج اجرای آن را ملاحظه کردیم.

فصل چهارم

روایی افراد شبکه‌ای از دورین‌ها باید مجزا

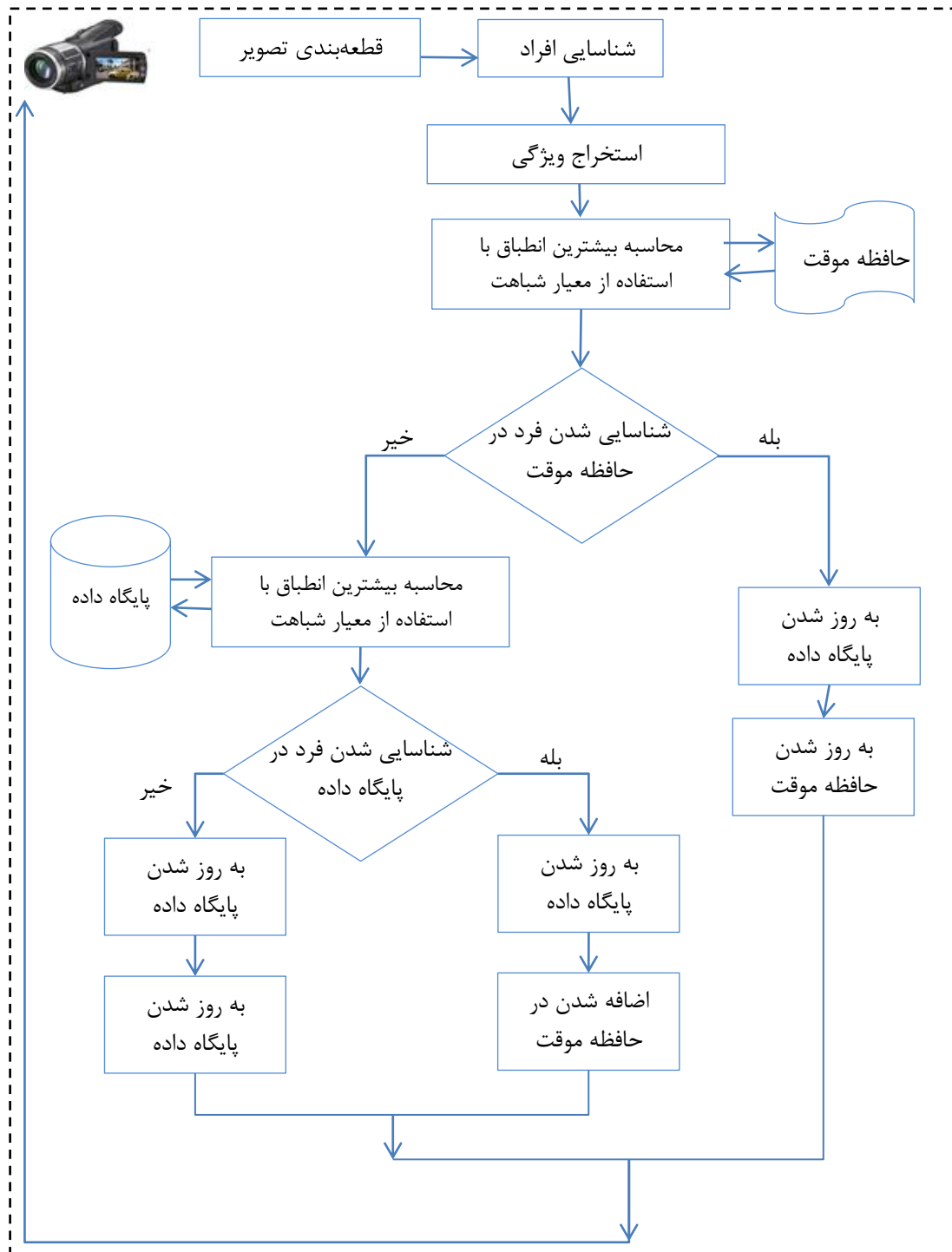
۴-۱- مقدمه

ردیابی افراد به معنی استخراج ویژگی‌های یک شی در یک فریم و تعقیب شی در فریم‌های بعدی بر اساس ویژگی‌های استخراج شده می‌باشد. به همین منظور، در این فصل به معرفی روال کلی برای ردیابی انسان در یک دوربین و شبکه‌ای از دوربین‌ها با دید مجزا (گسسته) پرداخته شده است. در همچنین سیستمی ردیابی به دو بخش اصلی شناسایی و ردیابی افراد در زاویه دید یک دوربین و شناسایی مجدد افراد در بین دوربین‌ها تقسیم می‌شود. که تمام مراحل لازم برای طراحی یک سیستمی که قادر باشد در همچین محیط‌های عمل ردیابی را انجام دهد در این فصل معرفی شده است. بعد از آنکه افراد با استفاده از تفاضل پس‌زمینه به خوبی از تصاویر استخراج شدند باید ویژگی‌هایی به منظور ردیابی افراد انتخاب شوند با توجه به مزایای که ویژگی‌های مبتنی بر ظاهر افراد در همچین سیستم‌ها داشته‌اند از دو ویژگی هیستوگرام رنگ و روش تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی^۱ برای شناسایی و ردیابی افراد استفاده شده است. استفاده از ویژگی هیستوگرام رنگ به دلیل سرعت (بار محاسباتی) و دقت انتخاب شد که برای از بین بردن تغییرات شدت روشنایی افراد در بین دوربین‌ها از تابع شدت روشنایی تجمعی استفاده شده و به منظور بالا بردن عملکرد ویژگی رنگ، بدن افراد با استفاده از فاصله نسبی قسمت‌ها مختلف بدن، به سه قسمت سر، نیم‌تنه و پایین‌تنه تقسیم شد. در این پایان‌نامه پیشنهاد شد که استفاده هریک از ویژگی‌ها با توجه به موقعیت افراد نسبت به محیط استفاده شوند به همین منظور با استفاده از یک سیستم فازی محیط تحت مراقبت توسط هر دوربین به شش قسمت متفاوت تقسیم شده است و این دو ویژگی با توجه به محل قرار گرفتن افراد استفاده شده است.

¹ Nonlinear Fuzzy Robust Principle Component Analysis

۴-۲- ردیابی و شناسایی مجدد افراد در چند دوربین

بعد استخراج افراد از تصاویر که در فصل قبل مورد بررسی قرار گرفت حالت نوبت شناسایی و ردیابی افراد می‌رسد. در این مرحله ابتدا باید مورد بررسی قرار گیرد که فردی استخراج شده آیا در فریم‌های قبلی وجود داشته یا نه. برای این منظور ابتدا ویژگی‌ها از افراد استخراج شده و سپس فرد با افراد موجود در حافظه موقت یا همان فریم‌های قبل مورد ارزیابی قرار می‌گیرد در صورتی که فرد در حافظه موقت وجود داشت پایگاه‌داده‌ای افراد و حافظه موقت به روز رسانی می‌شوند. ولی در صورتی که فرد جدید در حافظه موقت یافت نشده باشد پس فرد به عنوان شخص جدیدی می‌باشد که تازه وارد زاویه دید دوربین شده است. دو احتمال اینجا به وجود می‌آید که فرد تا کنون از زاویه دید دوربین عبور نکرده و یا اینکه فرد قبلاً در زاویه دید این دوربین یا دوربین‌ها دیگر بوده و دوباره وارد زاویه دید دوربین شده است. برای تشخیص این نکته ویژگی‌های استخراج شده از فرد با افراد موجود در پایگاه‌داده مورد ارزیابی قرار می‌گیرد در صورتی که فرد در پایگاه داده‌ها وجود داشته باشد فرد در حافظه موقت اضافه شده و پایگاه‌داده‌ها به روز رسانی می‌شود ولی در صورتی که این فرد در پایگاه‌داده هم وجود نداشته باشد به عنوان فردی معرفی می‌شود که تا کنون در زاویه دید دوربین‌ها دیده نشده است. در این حالت فرد جدید هم پایگاه‌داده و حافظه موقت اضافه می‌شود. مراحل کار ردیابی افراد در شبکه‌ای از دوربین‌ها به صورت شماتیک در شکل (۴-۱) نشان داده شده است.



شکل (۴-۱) شمای کلی از مراحل ردیابی و شناسایی مجدد انسان

۴-۳- بهبود عملکرد سیستم ردیابی افراد با روش پیشنهادی

در این پایان‌نامه برای بالا بردن عملکرد سیستم ردیابی در ابتدا محیط تحت مراقبت توسط هر دوربین

با استفاده از روش فازی که در قسمت بعد، توضیح داده می‌شود به شش قسمت نزدیک و روبه دوربین، متوسط روبه دوربین، دور رو به دوربین، نزدیک پشت به دوربین، متوسط پشت به دوربین و دور پشت به دوربین تقسیم می‌نماییم. در ادامه به معرفی ویژگی‌های معرفی شده در این پایان‌نامه، چگونگی استفاده از این ویژگی‌ها در مکان‌ها متفاوت از محیط مورد نظارت پرداخته شده است. در انتها مراحل کلی که در شکل (۴-۱) نشان داده شد با توجه به سیستم پیشنهادی تصحیح شده و بهبود داده شده است.

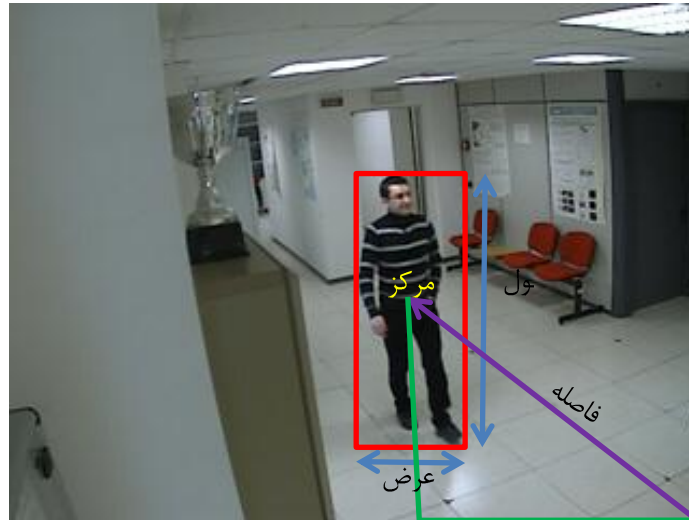
۴-۳-۱- موقعیت یا وضعیت افراد نسبت به دوربین‌ها

همانطور که در فصل‌های قبلی توضیح داده شده یکی از مشکلات استفاده از ویژگی برای شناسایی افراد، تغییر اندازه افراد در زاویه دید دوربین‌ها در واقع کوچک و بزرگ شدن ظاهر افراد در زمانی که فاصله آنها نسبت به دوربین‌ها تغییر می‌کند، می‌باشد، اگر بتوان اطلاعاتی از موقعیت افراد نسبت به دوربین‌ها بدست آورد و از ویژگی‌های هر قسمت برای آن قسمت به صورت جدا استفاده نمود این عمل می‌تواند کمک بسیاری برای شناسایی افراد انجام دهد به همین منظور در این پایان‌نامه با معرفی روشی جدید به تخمین موقعیت افراد نسبت به دوربین با استفاده از قسمت‌بندی صحنه‌ای که دوربین‌ها تحت پوشش قرار می‌دهند، می‌پردازیم. برای این منظور ما از جدول وضعیت که در بخش بعدی توضیح داده خواهد شد استفاده می‌نماییم. در این پایان‌نامه، برای اینکه بتوان موقعیت افراد نسبت به دوربین‌ها را بدست آورد به اطلاعات به مانند ارتفاع، طول، اندازه، زاویه نسبت به دوربین و فاصله پیکسلی افراد نیاز داریم که در ادامه به معرفی این ویژگی‌ها می‌پردازیم.

• ارتفاع و عرض

یکی از ویژگی‌های که در این پروژه برای بدست آوردن موقعیت افراد نسبت به دوربین استفاده شده فاصله بین ابتدا سر تا انتهای پا، به عبارتی قد افراد یا ارتفاع افراد می‌باشد. ویژگی دیگری که به ظاهر و موقعیت افراد نسبت به دوربین بستگی دارد پهنای بدن انسان یا همان عرض افراد می‌باشد که در

شکل (۲-۴) نمونه‌ای از این ویژگی‌ها را نشان می‌دهد. البته ویژگی عرض یا پهنا با توجه به زاویه دید فرد نسبت به دوربین می‌تواند تا حدودی متغیر باشد.



شکل (۲-۴) نمایشی از ویژگی‌های طول، عرض و فاصله برای یک فرد

• فاصله پیکسلی

ویژگی فاصله پیکسلی می‌تواند به عنوان یک پارامتر بسیار مؤثر برای بدست آوردن موقعیت افراد در تصاویر ویدیویی در زمانی که فاصله افراد نسبت به دوربین‌ها تغییر می‌کند در نظر گرفته شود. که برای بدست آوردن این فاصله نیاز به اطلاعاتی اضافه‌ای مثل محل قرار گرفتن دوربین نسبت به محیط می‌باشد یا یک نقطه ثابت برای هر دوربین را در نظر گرفته شود. با توجه به این نکته که این پروژه در محیط بسته انجام شده می‌توان محل قرار گرفتن دوربین‌ها را تا حدودی مشخص نمود. برای بدست آوردن فاصله پیکسلی افراد نسبت به دوربین‌ها می‌توان از معیارهای زیادی استفاده کرد، که در اینجا از فاصله اقلیدسی استفاده شده است. فاصله محل مورد نظر دوربین (نقطه ثابت)، نسبت به مرکز شی استخراج شده با استفاده از رابطه (۱-۴) بدست آورده می‌شود. شکل (۲-۴) یک نمونه از ویژگی‌های ارتفاع، عرض و فاصله برای یک فرد را نشان می‌دهد.

$$dist = \sqrt{(x_2 - Cx_1)^2 + (y_2 - Cy_1)^2} \quad (۱-۴)$$

در رابطه (۱-۴) Cx_1 و Cy_1 مختصات مرکز فرد مورد نظر برای شناسایی و ردیابی می‌باشد.

- اندازه یا حجم

برای بدست آوردن حجم یا اندازه افراد از تصویر سایه‌نمای استخراج شده از افراد استفاده می‌شود. در شکل (۳-۴) نمونه‌های از تصاویر سایه‌نمای چند فرد را نشان می‌دهد. با استفاده از رابطه (۲-۴) اندازه هر فرد استخراج شده محاسبه می‌شود. در رابطه زیر n و m ابعاد فرد می‌باشد.

$$Z = \sum_{i=1}^n \sum_{j=1}^m B(i, j) \quad (۲-۴)$$



شکل (۳-۴) نمایشی از افراد استخراج شده به همراه تصویر سایه‌نما

- زاویه

زاویه قرار گرفتن افراد نسبت به دوربین می‌تواند در شناسایی و ردیابی بسیار تاثیرگذار باشد بخصوص زمانی که از ویژگی‌های بیومتریک مثل صورت افراد یا طریقه راه رفتن برای ردیابی استفاده می‌شود. تشخیص این نکته که شخص روبه دوربین یا پشت به دوربین می‌باشد می‌تواند برای استخراج ویژگی بسیار حائز اهمیت باشد. برای این منظور ابتدا بردار حرکت هر پیکسل فرد نسبت به فریم‌های قبلی محاسبه می‌شود و با بدست آوردن برآیند بردارها، زاویه بردار حاصل نسبت به مبدا در نظر گرفته می‌شود. در واقع همه بردارها هم مبدا می‌شوند. برای بدست آوردن بردار حرکت از روش جریان نوری استفاده شده است. جریان نوری با دو روش معروف Horn-Schunck و Lucas-Kanade شناخته می‌شود که در این پروژه از روش Horn-Schunck استفاده شده است. که به صورت زیر محاسبه می‌شود. برای محاسبه جریان نوری در دو تصویر رابطه (۳-۴) را باید محاسبه نمود.

$$I_x u + I_y v + I_t = 0 \quad (۳-۴)$$

در رابطه بالا I_x ، I_y و I_t مشتق روشنایی و u ، v به ترتیب جریان نوری افقی و عمودی می‌باشد. در روش Horn-Schunck برای تخمین سرعت u ، v رابطه (۴-۴) را باید مینیمم نمود.

$$E = \iint (I_x u + I_y v + I_t)^2 dx dy + \alpha \iint \left\{ \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 \right\} dx dy \quad (۴-۴)$$

بردار حرکت در مرحله بعدی با استفاده از رابطه (۵-۴) محاسبه می‌شود.

$$u_{x,y}^{k+1} = \bar{u}_{x,y}^k - \frac{I_x [I_x \bar{u}_{x,y}^k + I_y \bar{v}_{x,y}^k + I_t]}{\alpha^2 + I_x^2 + I_y^2} \quad (۵-۴)$$

$$v_{x,y}^{k+1} = \bar{v}_{x,y}^k - \frac{I_y [I_x \bar{u}_{x,y}^k + I_y \bar{v}_{x,y}^k + I_t]}{\alpha^2 + I_x^2 + I_y^2} \quad (۶-۴)$$

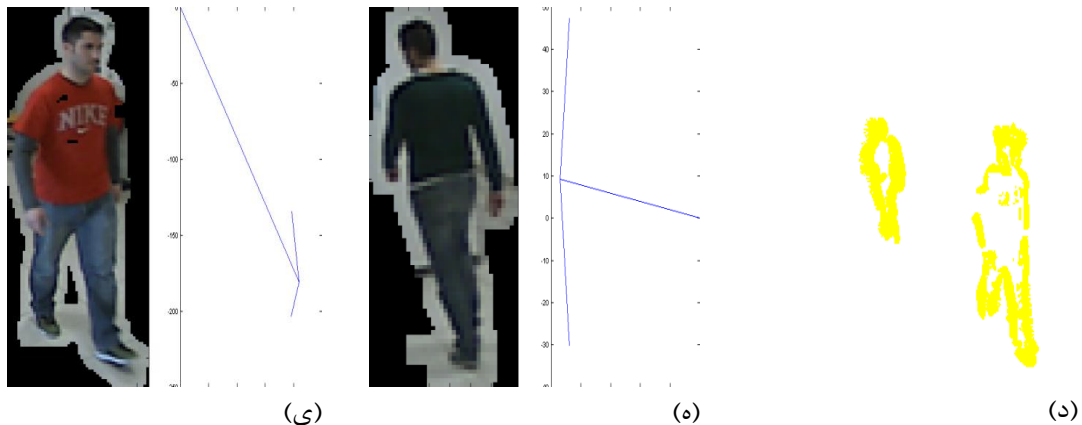
در این رابطه، $[u_{x,y}^k, v_{x,y}^k]$ سرعت تخمینی برای پیکسل (x,y) می‌باشد و $[\bar{u}_{x,y}^k, \bar{v}_{x,y}^k]$ میانگین همسایه‌های این پیکسل می‌باشد. برای $k=0$ سرعت اولیه را صفر در نظر گرفته شده است.

در رابطه (۳-۴)، I_x ، I_y با کرنل سوبل^۱ $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$ و I_t بین دو تصویر ۱ و ۲ از کرنل $\begin{bmatrix} -1 & 1 \end{bmatrix}$ استفاده شده است.

شکل (۴-۴) نمونه‌ای از جریان نوری بدست آمده با استفاده از روش Horn-Schunck را نشان می‌دهد.



¹ Sobel



شکل (۴-۴) بردار حرکت بدست آمده با روش Horn-Schunck. (الف) و (ب) دو فریم پشت سرهم از ویدیو (فریم-های ۶۲۲-۶۲۳). (ج) بردارهای حرکت در فریم ۶۲۳ (بردارهای حرکتی که از یک حدآستانه کمتر بودن را نمایش نداده-ایم). (د) بردار حرکت بدون پس‌زمینه. (ه) و (ی) جهت حرکت افراد بعد از بدست آوردن برآیند بردار حرکت به هر فرد به صورت جدا را نمایش می‌دهد.

در این روش در صورتی که برآیند بردارهای بدست آمده از افراد بزرگتر از صفر باشد نشان دهنده روبه دوربین (بازه حدودا بین ۹۰-۲۷۰) و در غیر این صورت پشت به دوربین (بازه حدودا بین ۰-۹۰ و ۲۷۰-۳۶۰) می‌باشد.

۴-۳-۲- جدول وضعیت افراد

در این پروژه می‌خواهیم موقعیت و محل قرار گرفتن افراد نسبت به دوربین را محاسبه نماییم برای همین منظور از جدول (۴-۱) برای بدست آوردن موقعیت یا وضعیت افراد استفاده شده است. در این جدول از پنج ویژگی اندازه، طول، عرض، زاویه و فاصله پیکسلی که در قسمت‌های قبل توضیح داده شده، استفاده کرده‌ایم. که ویژگی‌های معرفی شده به اندازه‌های کوچک، متوسط و بزرگ تقسیم می‌شوند با توجه به این که محیط بسته می‌باشد و مشخصات فیزیکی محیط تا حدودی قابل تشخیص می‌باشد. مقدار این ویژگی‌ها را برای افراد متفاوت می‌تواند بدست آورد به همین منظور در زمان راه-اندازی پارامترهای مورد نیاز برای کامل کردن این جدول محاسبه می‌شود.

جدول (۴-۱) جدول وضعیت افراد نسبت به دوربین‌ها

وضعیت	اندازه	ارتفاع	عرض	فاصله پیکسلی	زاویه
وضعیت ۱	بزرگ	بزرگ	کوچک	بزرگ	منفی
وضعیت ۲	بزرگ	متوسط	متوسط	بزرگ	منفی
وضعیت ۳	متوسط	کوچک	بزرگ	متوسط	منفی
وضعیت ۴	بزرگ	بزرگ	کوچک	بزرگ	مثبت
وضعیت ۵	بزرگ	متوسط	متوسط	بزرگ	مثبت
وضعیت ۶	متوسط	کوچک	بزرگ	متوسط	مثبت

در این پروژه برای بدست آوردن مقادیر این ویژگی‌های بعد از نصب دوربین در مرحله راه‌اندازی^۱، یک فرد در محیط حرکت کرده مقادیر بیشینه و کمینه هر ویژگی را برای آن محاسبه می‌نماییم و با توجه به این مقادیر پارامترها برای جدول وضعیت بدست آورده می‌شود.

برای بدست آوردن مقدار ویژگی اندازه که با پارامترهای کوچک، متوسط و بزرگ مشخص می‌شود از رابطه‌های زیر استفاده شده است.

$$big_size > Min_size + \frac{2}{3}(Max_size - Min_size) \pm \varepsilon \quad (۷-۴)$$

$$Min_size + \frac{1}{3}(Max_size - Min_size) \pm \varepsilon < Ave_size... \quad (۸-۴)$$

$$< Min_size + \frac{2}{3}(Max_size - Min_size) \pm \varepsilon$$

$$small_size < Min_size + \frac{1}{3}(Max_size - Min_size) \pm \varepsilon \quad (۹-۴)$$

برای بدست آوردن مقدار ویژگی ارتفاع از رابطه‌های زیر استفاده شده است.

$$big_Height > Min_Height + \frac{2}{3}(Max_Height - Min_Height) \pm \varepsilon \quad (۱۰-۴)$$

^۱ Setup

$$Min_Height + \frac{1}{3}(Max_Height - Min_Height) \pm \varepsilon < Ave_Height... \quad (11-4)$$

$$< Min_Height + \frac{2}{3}(Max_Height - Min_Height) \pm \varepsilon$$

$$small_Height < Min_Height + \frac{1}{3}(Max_Height - Min_Height) \pm \varepsilon \quad (12-4)$$

برای بدست آوردن مقدار عرض افراد که با پارامترهای بزرگ، متوسط و کوچک از رابطه‌های زیر استفاده شده است.

$$big_Width > Min_Width + \frac{2}{3}(Max_Width - Min_Width) \pm \varepsilon \quad (13-4)$$

$$Min_Width + \frac{1}{3}(Max_Width - Min_Width) \pm \varepsilon < Ave_Width... \quad (14-4)$$

$$< Min_Width + \frac{2}{3}(Max_Width - Min_Width) \pm \varepsilon$$

$$small_Width < Min_Width + \frac{1}{3}(Max_Width - Min_Width) \pm \varepsilon \quad (15-4)$$

برای بدست آوردن مقدار فاصله که با پارامترهای بزرگ و کوچک مشخص می‌شود از رابطه‌های زیر استفاده شده است.

$$big_Dist > Min_Dist + \frac{2}{3}(Max_Dist - Min_Dist) \pm \varepsilon \quad (16-4)$$

$$Min_Dist + \frac{1}{3}(Max_Dist - Min_Dist) \pm \varepsilon < Avg_Dist... \quad (17-4)$$

$$< Min_Dist + \frac{2}{3}(Max_Dist - Min_Dist) \pm \varepsilon$$

$$small_Dist < Min_Dist + \frac{1}{3}(Max_Dist - Min_Dist) \pm \varepsilon \quad (18-4)$$

وضعیت افراد را می‌توان با استفاده از ویژگی‌های معرفی شده نسبت به دوربین به شش قسمت تقسیم نمود. این قوانین را می‌توان به صورت یک پایگاه داده از قوانین فازی برای بدست آوردن موقعیت افراد در نظر گرفت. به همین منظور در این پایان‌نامه برای بدست آوردن موقعیت افراد از سیستم فازی استفاده شده است.

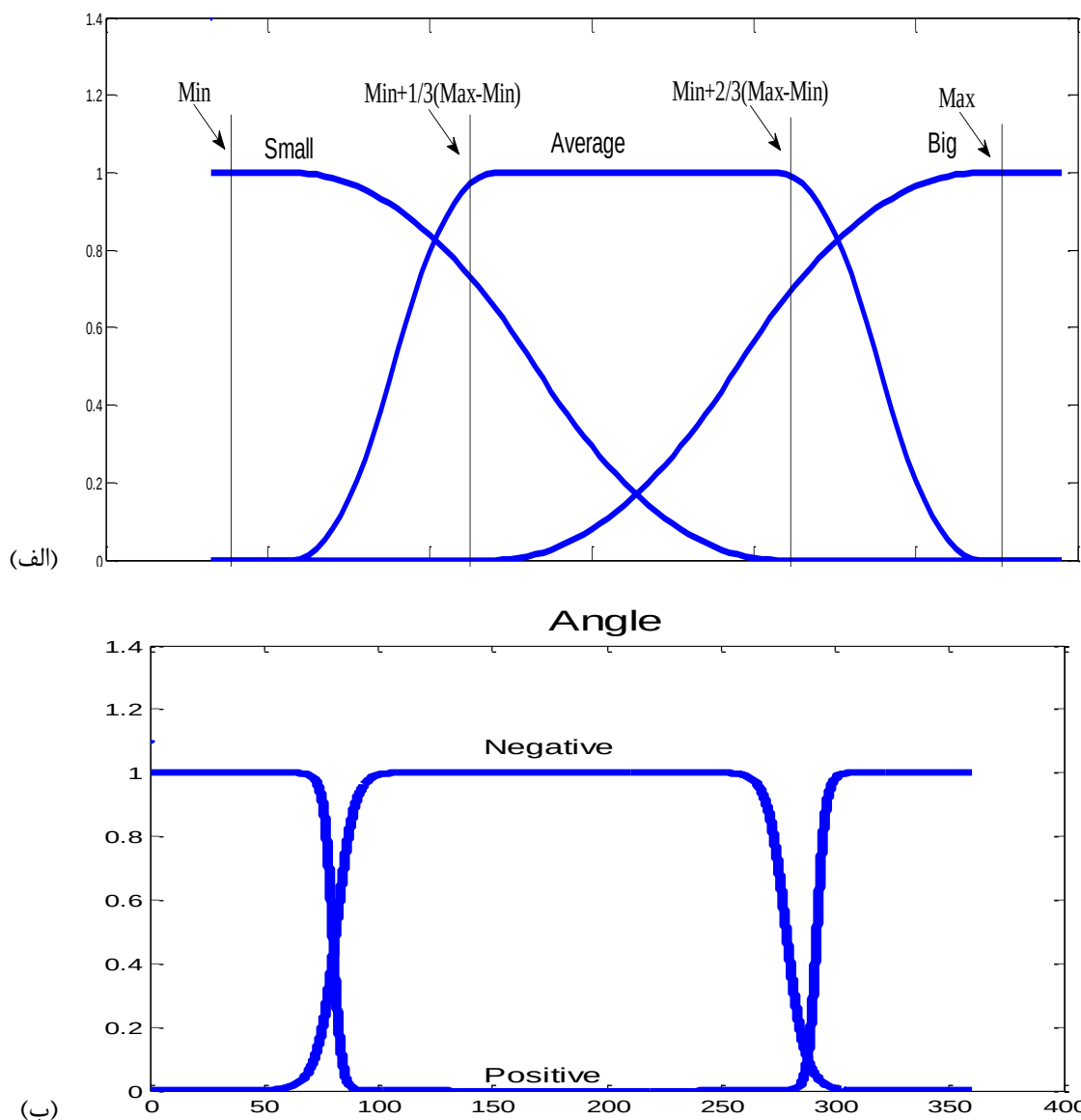
جدول (۲-۴) وضعیت افراد نسبت به دوربین‌ها با در نظر گرفتن ویژگی‌های و مقادیری که برای هر ویژگی معرفی شده نشان می‌دهد.

جدول (۲-۴) جدول وضعیت و مقادیر ویژگی‌های معرفی شده

وضعیت	ارتفاع	عرض	فاصله پیکسلی	اندازه	زاویه
وضعیت ۱	Big	Big	Small	Big	Negative
وضعیت ۲	Big	Avg	Avg	Big	Negative
وضعیت ۳	Avg	Small	Big	Avg	Negative
وضعیت ۴	Big	Big	Small	Big	Positive
وضعیت ۵	Big	Avg	Avg	Big	Positive
وضعیت ۶	Avg	Small	Big	Avg	Positive

۳-۳-۴- استفاده از روش فازی برای بدست آوردن وضعیت افراد نسبت به دوربین‌ها

در این سیستم فازی ورودی‌ها، پنج ویژگی معرفی شده در قسمت قبل می‌باشد و خروجی این سیستم یکی از شش وضعیت افراد نسبت به دوربین می‌باشد. هریک از این وضعیت‌ها در حالت‌های مختلف می‌توانند با یکدیگر هم‌پوشانی دارد. ابتدا باید قواعد حاکم بر این سیستم را بدست آورد برای این منظور از پایگاه معرفت که در جدول (۲-۴) وجود دارد استفاده می‌نماییم. با توجه به این جدول هر یک از ویژگی‌ها را به صورت مجموعه‌ای از عبارات فازی به صورت Small, Average و Big در نظر گرفت که در شکل (۴-۵) این عبارات با استفاده از توابع فازی نشان داده شده است.



شکل (۴-۵) مقدار هر ویژگی جدول (۴-۲) به عنوان ورودی به سیستم فازی. الف) تابع فازی برای ویژگی‌های ارتفاع، عرض، اندازه و فاصله پیکسلی. ب) تابع فازی برای زاویه

البته ویژگی زاویه به صورت شکل (۴-۵ ب) می‌باشد چون در صورتی که زاویه برآیند بردارها افراد مثبت، فرد رو به دوربین و در غیر این صورت پشت به دوربین می‌باشد.

بعد از بدست آوردن متغیرهای ورودی سیستم فازی نوبت به قواعد مورد استفاده از سیستم می‌رسد که در اینجا از شش قاعده استفاده شده که از جدول (۴-۲) بدست آمده و به صورت زیر می‌باشند.

قاعده ۱: اگر ارتفاع فرد بزرگ (Big) و مقدار عرض آن بزرگ (Big) و مقدار فاصله آن کوچک (Small) مقدار حجم آن بزرگ (Big) و زاویه پشت به دوربین (Negative) وضعیت ۱ می‌باشد.

قاعده ۲: اگر ارتفاع فرد بزرگ (Big) و مقدار عرض آن متوسط (Avg) و مقدار فاصله آن متوسط (Avg) و مقدار حجم آن بزرگ (Big) و زاویه پشت به دوربین (Negative) وضعیت ۲ می‌باشد.

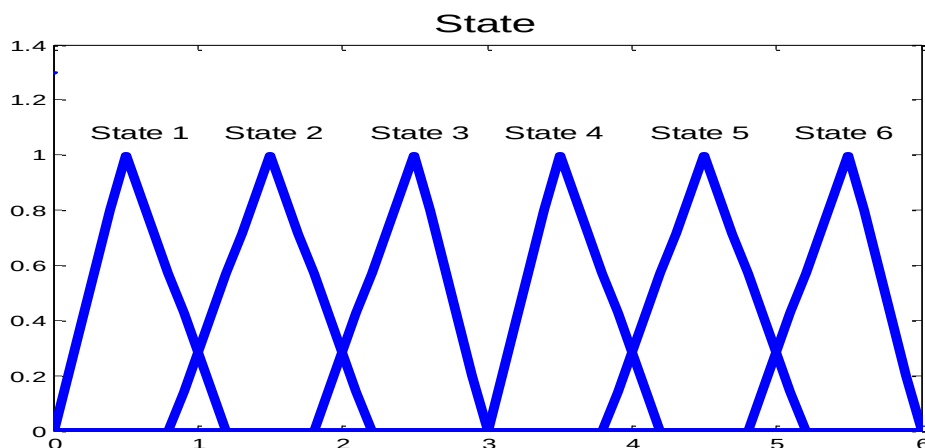
قاعده ۳: اگر مقدار ارتفاع فرد متوسط (Avg) و مقدار عرض آن کوچک (Small) و مقدار فاصله آن بزرگ (Big) و مقدار حجم آن متوسط (Avg) و زاویه پشت به دوربین (Negative) وضعیت ۳ می‌باشد.

قاعده ۴: اگر ارتفاع فرد بزرگ (Big) و مقدار عرض آن بزرگ (Big) و مقدار فاصله آن کوچک (Small) و مقدار حجم آن بزرگ (Big) و زاویه رو به دوربین (Positive) وضعیت ۴ می‌باشد.

قاعده ۵: اگر ارتفاع فرد بزرگ (Big) و مقدار عرض آن متوسط (Avg) و مقدار فاصله آن متوسط (Avg) و مقدار حجم آن بزرگ (Big) و زاویه رو به دوربین (Positive) وضعیت ۵ می‌باشد.

قاعده ۶: اگر مقدار ارتفاع فرد متوسط (Avg) و مقدار عرض آن کوچک (Small) و مقدار فاصله آن بزرگ (Big) و مقدار حجم آن متوسط (Avg) و زاویه رو به دوربین (Positive) وضعیت ۶ می‌باشد.

در شکل (۴-۶) تابع عضویت مربوط به مقادیر زبانی متغیر خروجی نشان داده شده است.



شکل (۴-۶) وضعیت افراد نسبت به دوربین به عنوان خروجی سیستم فازی

پس از بدست آوردن قواعد حاکم بر سیستم و تشکیل پایگاه معرفت، نیاز به موتور استنتاج داریم تا با پذیرفتن ورودی‌های فازی براساس قواعد پایگاه معرفت، خروجی فازی مناسب را ایجاد نماییم. این

امکان وجود دارد که قواعد فازی با یکدیگر هم‌پوشانی داشته باشند. به عبارت دیگر، تابع عضویت متغیر ورودی با دو یا سه تابع عضویت از عبارت متغیر زبانی تطابق نسبی داشته باشد که در این نتیجه با دو یا سه قاعده از قواعد پایگاه معرفت، قابل استنتاج خواهد شد. در این حالت باید یک استراتژی رقابتی برای استنتاج وجود داشته باشد تا بتوان خروجی فازی مناسب را ایجاد کرد که برای این منظور دو روش مدل ممدانی و سانگو به عنوان معروف‌ترین روش‌ها مطرح می‌باشد که در این پایان نامه با توجه به فرم قواعد ممدانی که در رابطه (۴-۱۹) نمونه‌ای از آن را نشان می‌دهد در این پروژه استفاده شده است.

$$\text{if } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ then } z \text{ is } C_1 \quad (۴-۱۹)$$

تا حالا مقادیر بدست آمده از روش مقادیر فازی می‌باشند و لازم است تا این مقادیر یا مجموعه‌های فازی را به مقادیر غیرفازی^۱ مناسب تبدیل نماییم. روش‌های زیادی برای غیرفازی کردن وجود دارد که از مهمترین این روش‌ها می‌توان به مرکز ثقل^۲، مرکز مجموع‌ها^۳، ارتفاع، بزرگترین سطح، متوسط ماکزیمم^۴ و ... اشاره نمود. در اینجا از روش ارتفاع به دلیل سادگی و سرعت آن استفاده نموده‌ایم. در ادامه به توضیح این روش می‌پردازیم.

۴-۳-۴- غیر فازی کردن به روش ارتفاع

اگر هر قاعده، خروجی C_i را داشته که ارتفاع آن (α_i) قدرت آن قاعده در رقابت را مشخص می‌نماید و مقدار قله C_i برابر مقداری از Z است که در آن، بیشترین مقدار را داراست که با $z^{(i)}$ نمایش داده می‌شود. در این روش برای بدست آوردن خروجی قطعی از رابطه (۴-۲۰) استفاده می‌شود.

$$Z^* = \frac{\sum_{k=1}^n \alpha_k \cdot z_k}{\sum_{k=1}^n \alpha_k} \quad (۴-۲۰)$$

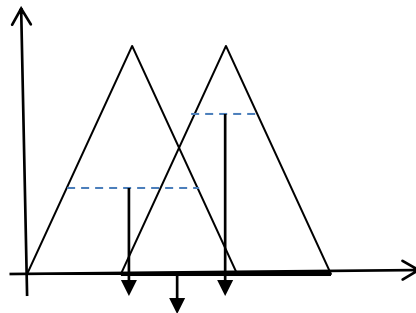
^۱ defuzzy

^۲ Center Of Gravity (COG)=Center Of Area (COA)

^۳ Center Of Sums(COS)

^۴ Mean Of Maxima(MOM)

شکل (۷-۴) یک نمونه از غیرفازی کردن به روش ارتفاع را نشان می‌دهد.



شکل (۷-۴) غیرفازی کننده به روش ارتفاع

۴-۴- استخراج ویژگی برای شناسایی افراد

بعد مرحله شناسایی وضعیت یا موقعیت افراد نسبت به دوربین‌ها نوبت آن می‌رسد که با استفاده از روش‌های موجود، ابتدا به شناسایی افراد و در مرحله بعد ردیابی آنها بپردازیم. همانطور که در فصل مروری بر کارهای گذشته ذکر شد روش‌های زیادی برای شناسایی و ردیابی افراد در شبکه‌ای از دوربین‌ها وجود دارد. برای ردیابی افراد در محیط‌هایی که از شبکه‌ای از دوربین‌ها با دید مجزا استفاده می‌شود ویژگی حرکت به مانند مکان، سرعت^۱ و شتاب نمی‌تواند برای شناسایی مجدد افراد مفید باشد. زیرا در این نوع از محیط‌ها نقاط ورود و خروج زیادی وجود دارند این در حالی است که ویژگی حرکت به مکان شی بستگی دارد. این امکان وجود دارد که اشیاء از یک مکان خارج شوند و دوباره در همان صحنه، ولی این بار از مکان دیگر، وارد شوند. همانطور که شکل (۱-۱) مشاهده شد که حالت‌های مختلف از ورود و خروج افراد را در این نوع از سیستم‌ها نشان می‌دهد که با توجه به متغیر بودن زمان و مکان اشیاء در هنگام ورود و خروج، می‌توان نتیجه گرفت که روش‌های مبتنی بر زمان-مکان نمی‌توانند مناسب باشند.

به طور مشابه ذکر شد که خواص هندسی نمی‌توانند نتایج خوبی برای شناسایی دوباره اشیاء برای اشیاء غیرصلب به دلیل، کیفیت پایین تصاویر و تغییر اندازه بسیار زیاد (دور و نزدیک شدن شی نسبت

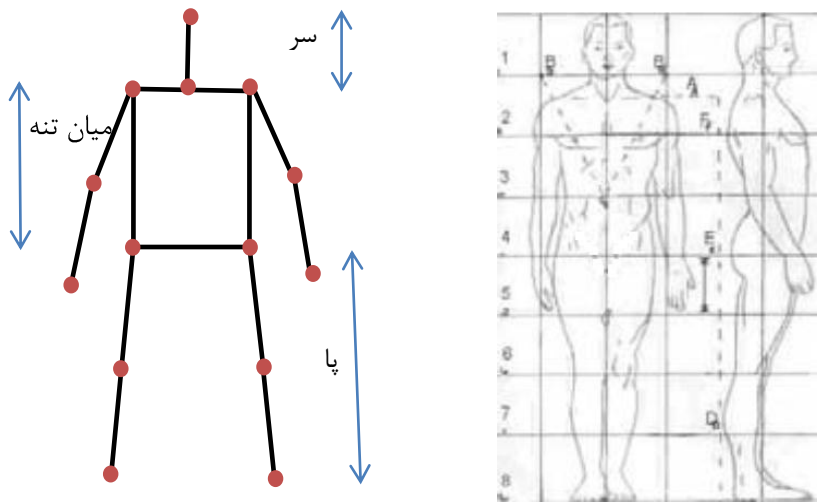
¹ Velocity

به دوربین‌های متفاوت)، از خود نشان دهند [۶۴]. دسته دیگری از ویژگی‌ها که در ردیابی مورد استفاده قرار گرفته شده‌اند اطلاعات بیومتریکی افراد مثل چهره و طریقه راه رفتن افراد می‌باشند این نوع از ویژگی‌ها نیز نمی‌توانند در ردیابی در شبکه‌ای از دوربین‌ها مناسب باشند زیرا استفاده از این نوع از ویژگی‌ها مثلا چهره ملزم به حرکت افراد در فاصله کم نسبت به دوربین می‌باشد و از طرفی برای تشخیص صحیح چهره نیاز به تصاویری با کیفیت تقریبا بالا می‌باشد و همچنین در هنگام استفاده از الگوریتم‌های مربوط به نحوه راه رفتن نیاز به پردازش‌های سنگین می‌باشد و نمونه‌های ویدیویی باید کاملا کنترل شده، در یک محیط و توسط یک دوربین تهیه شده باشند تا برای ردیابی استفاده شوند با توجه به این محدودیت‌ها که در استفاده از ویژگی‌های بیومتریکی به منظور شناسایی افراد وجود دارد می‌توان این نتیجه را گرفت که این نوع از ویژگی‌ها در محیط‌های مراقبتی که از چندین دوربین با تصاویر کیفیت پایین و یا معمولی استفاده می‌نماییم برای شناسایی مجدد و ردیابی نمی‌تواند مفید باشند. به همین منظور ما به در این پایان‌نامه از ویژگی‌های مبتنی بر ظاهر افراد به مانند رنگ استفاده نموده‌ایم. که در ادامه به معرفی ویژگی‌های که در این پایان‌نامه استفاده شده می‌پردازیم.

۴-۴-۱- رنگ

یکی از ویژگی‌های ظاهری برای شناسایی افراد در شبکه‌ای از دوربین‌ها رنگ می‌باشد. می‌توان رنگ را به عنوان بارزترین ویژگی ظاهری در سیستم‌های ردیابی دانست که معمولا به شکل هیستوگرام رنگ مورد استفاده قرار می‌گیرد. هیستوگرام رنگ به دلیل اینکه، ویژگی‌ای بسیار ساده و از لحاظ بار محاسباتی هم بسیار به صرفه می‌باشد، نمی‌توان به راحتی از کنار این ویژگی با ارزش رد شد به همین منظور در این پایان‌نامه از این ویژگی با ارزش برای شناسایی و ردیابی افراد با توجه به وضعیت افراد که توسط جدول (۴-۱) بدست می‌آوریم استفاده خواهیم نمود. ولی استفاده از هیستوگرام به تنهایی نمی‌تواند برای شناسایی مجدد افراد نتایج رضایت بخشی از خود نشان دهد مثلا با توجه به شباهت

بسیار زیاد رنگ لباس قسمت نیم‌تنه یک فرد یا قسمت پایین‌تنه فرد دیگر، این امکان وجود دارد که افراد به اشتباه به عنوان یک فرد شناسایی شوند به همین منظور در این پایان‌نامه به معرفی روشی برای تقسیم‌بندی بدن انسان با توجه به اطلاعات کلینیکی خواهیم پرداخت که به صورت مشروح در قسمت بعدی توضیح داده می‌شود.



شکل (۴-۸) نسبت های اعضای و مدل مورد استفاده‌ی بدن انسان [۹۸].

۴-۴-۲- قسمت‌های مختلف بدن انسان

بدن انسان به شکل یک سری قطعات به هم متصل رفتار می‌کند. برای تقسیم‌بندی قسمت‌های مختلف بدن انسان نیاز به نسبت‌های نسبی اندازه قسمت‌های مختلف با توجه به کل قد انسان داریم شکل (۴-۹)، که برای این منظور از فواصل نسبی، مطالعات کلینیکی^۱ پایگاه داده‌های آنالیز حرکات انسان و مطالعات نسبت‌های موجود در بدن انسان، استفاده می‌کنیم [۹۸ و ۹۹]. در مجموع می‌توان بدن انسان را به ارتفاع هشت سر مدل کرد و نسبت هر یک از اعضاء را نسبت به قد بیان نمود. شکل (۴-۸) به خوبی مبین این مسئله است و جدول (۴-۳) این نسبت‌ها را بیان می‌کند. در این پایان‌نامه ما بدن انسان را توجه به نسبت‌های موجود در این جدول به سه قسمت بالاتنه (سر)، نیم‌تنه و پایین‌تنه (پاها) تقسیم می‌نماییم. شکل (۴-۹) نتایج حاصل از قسمت‌بندی بدن انسان برای دو فرد را نشان می‌دهد.

¹ Clinical Studies

همانطور که از شکل مشخص می‌باشد بدست آوردن قسمت‌های مختلف بدن انسان ارتباط چندانی با محل قرار گرفتن افراد در محیط و اندازه افراد ندارد یعنی در واقع افراد با هر اندازه و موقعیتی نسبت به دوربین باشند این قسمت‌ها به درستی از هم جدا می‌شوند.



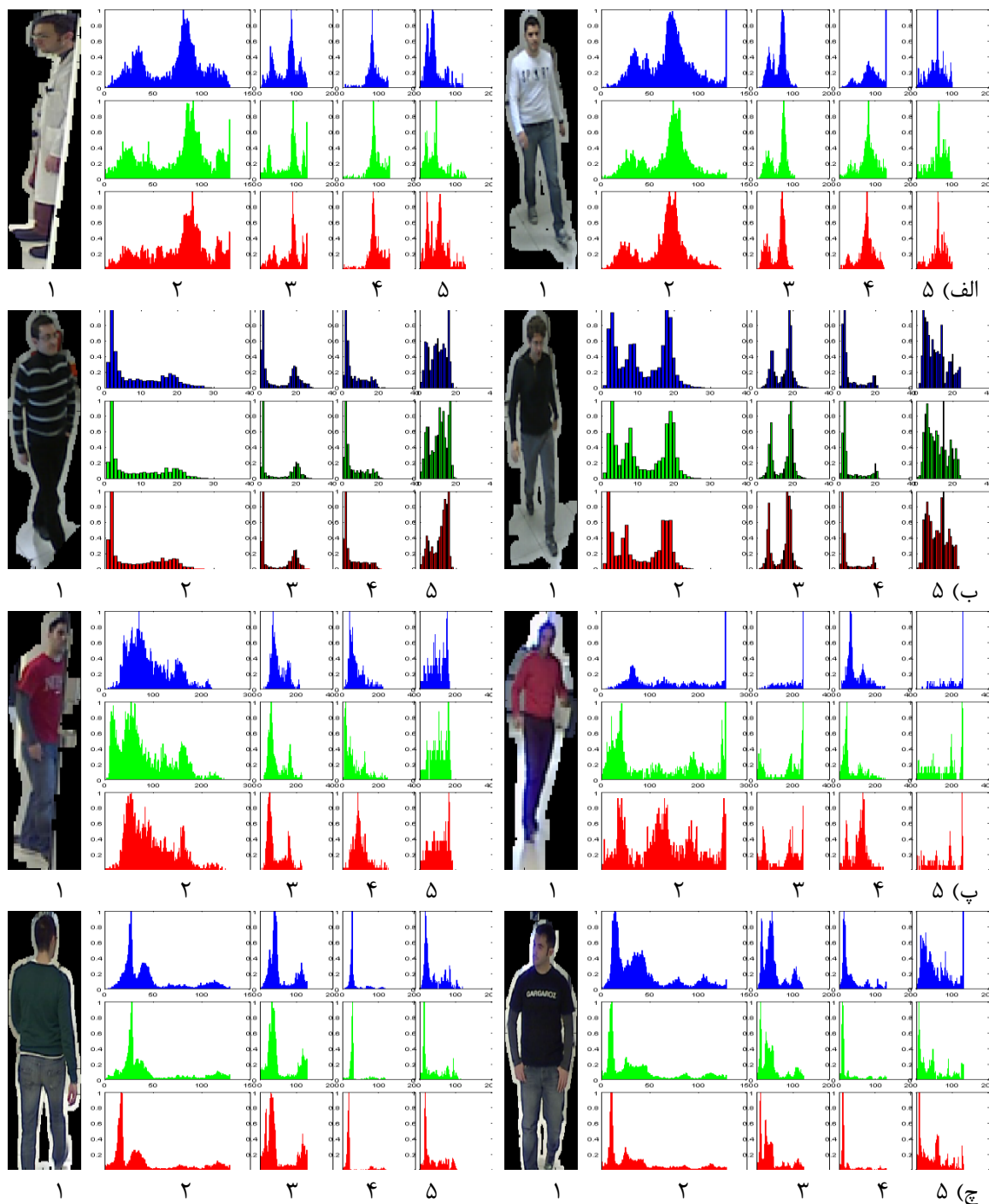
شکل (۴-۹). قسمت‌بندی بدن دو فرد با اندازه (قد) متفاوت با استفاده از روش پیشنهادی
جدول (۴-۳) اندازه نسبت‌های موجود در قسمت‌های مختلف نسبت به قد را نشان می‌دهد.

جدول (۴-۳) اندازه نسبی قسمت‌های مختلف	
طول نسبی (به واحد)	قطعه
۸	قد
۲	ساعد
۱,۵	بازو
۱,۲۵	گردن و سر
۱,۲۵	بدن (بالا تنه)
۲	فاصله بین دو شانه
۲	فاصله بین دو لگن
۲	ران
۲	ساق پا
۱	پا (از مچ به پایین)

۴-۳-۴- هیستوگرام قسمت‌های مختلف بدن انسان

در بسیاری از تحقیقات از هیستوگرام رنگ ظاهر جهت شناسایی افراد در دوربین‌های مختلف استفاده شده است [۴۲ و ۴۳]. در این مطالعات از هیستوگرام رنگ کل بدن انسان برای ردیابی بهره گرفته شده است. زمانی که هیستوگرام رنگ دو فرد، بدلیل شباهت رنگی بخش‌های متناظر به هم، نزدیک باشد می‌توان بجای کل بدن انسان به مقایسه قسمت‌های متناظر به صورت جداگانه پرداخت. به عنوان مثال، در شکل (۴-۱۱-الف) موردی را نشان می‌دهد که، هیستوگرام رنگی کل بدن دو شخص

(دو فرد متفاوت) به هم نزدیک می‌باشند ولی هیستوگرام رنگی پاهای این اشخاص تفاوت زیادی دارد، برای نشان دادن این مقایسه به صورت کمی از معیار شباهت Bahhatecheryya به عنوان معیار مقایسه استفاده می‌نماییم این معیار عددی در بازه صفر تا یک برمی‌گرداند که عدد یک نشان دهنده بیشترین شباهت بین افراد می‌باشد و عدد صفر حاکی از آن می‌باشد که این افراد هیچ شباهتی باهم ندارند، این معیار در بخش بعدی به طور کامل معرفی می‌شود. در همین راستا در جدول (۴-۴) (تصویر الف) این دو فرد با هم مقایسه شده‌اند که میزان شباهت هیستوگرام با در نظر گرفت کل بدن بیشتر از ۰,۶ می‌باشد که این مقدار نشان دهنده این نکته می‌باشد که هیستوگرام کل بدن این دو فرد، نزدیک به هم می‌باشد. اگر فقط این معیار در نظر گرفته شود باید این دو فرد را به عنوان یک فرد در هنگام شناسایی مجدد در نظر گرفت، ولی با دقت به هیستوگرام قسمت‌های مختلف تفاوت زیادی در میزان شباهت بین پایین‌تنه این دو فرد وجود دارد این تفاوت را می‌توان از جدول (۴-۴) (تصویر الف) به صورت کمی مشاهده نمود که میزان شباهت همه‌ی رنگ‌های پایین‌تنه کمتر از ۰,۵ می‌باشد، بنابراین می‌توان از این ویژگی جهت شناسایی و متمایز نمودن این دو فرد استفاده نمود. همچنین در شکل (۴-۱۱-ب) و شکل (۴-۱۱-پ)، موارد دیگری را نشان می‌دهد که به طور مشابه، رنگ میان‌تنه دو فرد بسیار به هم نزدیک می‌باشد (جدول (۴-۴) نشان دهنده این مطلب می‌باشد که میزان شباهت در هر دو مورد بیشتر از ۰,۶ می‌باشد) و از طرفی که نیم‌تنه حجم زیادی از بدن انسان را دارد، بنابراین اگر از هیستوگرام تمام بدن افراد استفاده شود این دو فرد به عنوان یک فرد در نظر گرفته می‌شود. ولی اگر از رنگ قسمت‌های متناظر برای بدست‌آوردن شباهت بین افراد استفاده شود رنگ پایین‌تنه این دو فرد را از هم متمایز می‌نماید که معیار در هر دو مورد کمتر از ۰,۵ می‌باشد. به همین منظور، در این پایان‌نامه با توجه به روش پیشنهاد شده، بدن انسان به سه قسمت تقسیم شده و هیستوگرام هر قسمت به صورت جدا محاسبه شده و بایکدیگر مقایسه می‌گردند.



شکل (۴-۱۰) هیستوگرام افراد متفاوت در فضای رنگی RGB و هیستوگرام هر قسمت از بدن به صورت جدا. در شکل هیستوگرام افراد با binهای متفاوت را نشان داده‌ایم، (الف و چ) ۱۲۸، (ب) ۳۲ و (پ) ۲۵۶. هر تصویر به چند قسمت تقسیم شده که در، قسمت (۱) فرد، قسمت (۲) هیستوگرام کل بدن برای رنگ‌ها آبی، سبز و قرمز، قسمت (۳) هیستوگرام پایین تنه (پا) برای رنگ‌ها آبی، سبز و قرمز، قسمت (۴) هیستوگرام میان تنه برای رنگ‌ها آبی، سبز و قرمز، قسمت (۵) هیستوگرام بالا تنه (سر) برای رنگ‌ها آبی، سبز و قرمز را نشان می‌دهد.

در شکل (۴-۱۱) نمونه دیگری را نشان داده‌ایم که این دو فرد از لحاظ هیستوگرام کل بدن و

پایین‌تنه بسیار به هم نزدیک می‌باشند ولی تفاوت قابل ملاحظه‌ای در میان‌تنه این افراد وجود دارد که سبب تمایز این اشخاص می‌شود. این تفاوت میان‌تنه و شباهت پایین‌تنه را می‌توان نیز به صورت کمی با استفاده از جدول (۴-۴) مشاهده کرد. جدول (۴-۴) میزان شباهت بین هر دو فرد هر قسمت شکل (۴-۱۱) را به صورت کمی با استفاده از معیار Bahhatecheryya نشان می‌دهد.

جدول (۴-۴) میزان شباهت هیستوگرام افراد شکل (۴-۱۱) با استفاده از معیار شباهت Bahhatecheryya

تصویر	Bin	رنگ قرمز کل بدن	رنگ سبز کل بدن	رنگ آبی کل بدن	رنگ قرمز میان تنه	رنگ سبز میان تنه	رنگ آبی میان تنه	رنگ قرمز پایین تنه	رنگ سبز پایین تنه	رنگ آبی پایین تنه
الف	۱۲۸	۰.۶۶	۰.۷	۰.۷۵	۰.۶۴	۰.۷۳	۰.۷	۰.۴۸	۰.۵	۰.۴۶
ب	۳۲	۰.۷۲	۰.۶۹	۰.۷۳	۰.۷۵	۰.۷۲	۰.۷۷	۰.۴۲	۰.۴۲	۰.۴۲
پ	۲۵۶	۰.۶۱	۰.۶۲	۰.۶۳	۰.۶	۰.۷	۰.۷۱	۰.۴۵	۰.۳۲	۰.۴۵
چ	۱۲۸	۰.۶۳	۰.۶۴	۰.۶۹	۰.۴۲	۰.۳۵	۰.۴۶	۰.۷۵	۰.۷۴	۰.۷۵

۴-۴-۴- تابع انتقال روشنایی تجمعی

همواره استفاده از رنگ یا مشتقات رنگی به دلیل حساسیت به شریط نوری و پارامترهای دوربین باعث ایجاد محدودیت‌های در تشخیص و ارتباط صحیح بین مشاهدات می‌شود [۱۰۰]. یکی از مشکلاتی در شناسایی افراد در شبکه‌ای از دوربین‌ها، تغییر شدت روشنایی در زمانی که افراد از یک دوربین به دوربین دیگر منتقل می‌شوند، می‌باشد که یکی از راه حل‌های ارائه شده برای حل این مشکل استفاده از تابع انتقال رنگ^۱ یا تابع انتقال روشنایی^۲ قبل از محاسبه‌ی میزان شباهت می‌باشد [۷۱] و [۱۰۱]. به همین منظور برای کم کردن تاثیر شدت روشنایی از یک تابع انتقال روشنایی که در مرجع‌های [۶۴] و [۱۰۲] پیشنهاد داده شده استفاده نموده‌ایم. در این مرجع از تکنیک هیستوگرام تجمعی معکوس^۳ استفاده نموده است. برای محاسبه هیستوگرام تجمعی از رابطه (۴-۲۶) استفاده می‌شود.

$$\hat{H}(B_m) = \sum_{k=1}^m \sum_{l=1}^N I(B_k) \quad (۴-۲۶)$$

^۱ Color Transfer Function

^۲ Brightness Transfer Function

^۳ Inverted Cumulative Histogram

در رابطه بالا، N تعداد نمونه‌های مجموعه آموزش را نشان داده و $B_1, \dots, B_m, \dots, B_M$ مقدار روشنایی از binهای هیستوگرام H_i و H_j می‌باشد. $\hat{H}_i$ هیستوگرام تجمعی برای H_i می‌باشد. برای پیدا کردن بهترین انطباق بین binهای هیستوگرام از تابع انتقال روشنایی (BTF) که در رابطه (۴-۲۷) نشان داده شده استفاده کرده‌ایم.

$$f_{i,j} = \hat{H}_j^{-1}(\hat{H}_i(B_m)) \quad (۴-۲۷)$$

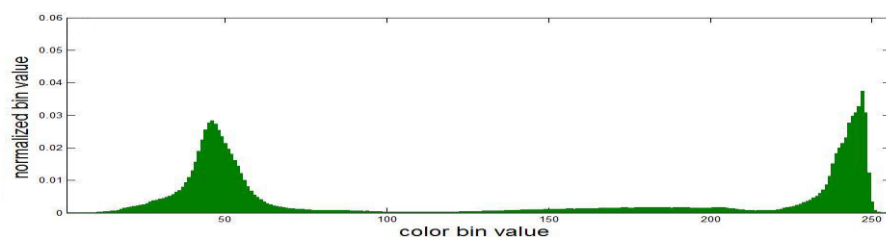
تابع $f_{i,j}(B_m)$ به صورت زیر بیان می‌شود:

- پیدا کردن بهترین انطباق بین bin، B_m از $\hat{H}_i$ با همه binهای هیستوگرام $\hat{H}_j$
- اگر bin، B_m از $\hat{H}_i$ با B_n از $\hat{H}_j$ انطباق داشت سپس $\hat{H}_i(B_{m+1})$ می‌تواند فقط با $\hat{H}_j(B_w)$ انطباق داشته باشد: ($w \geq n$)

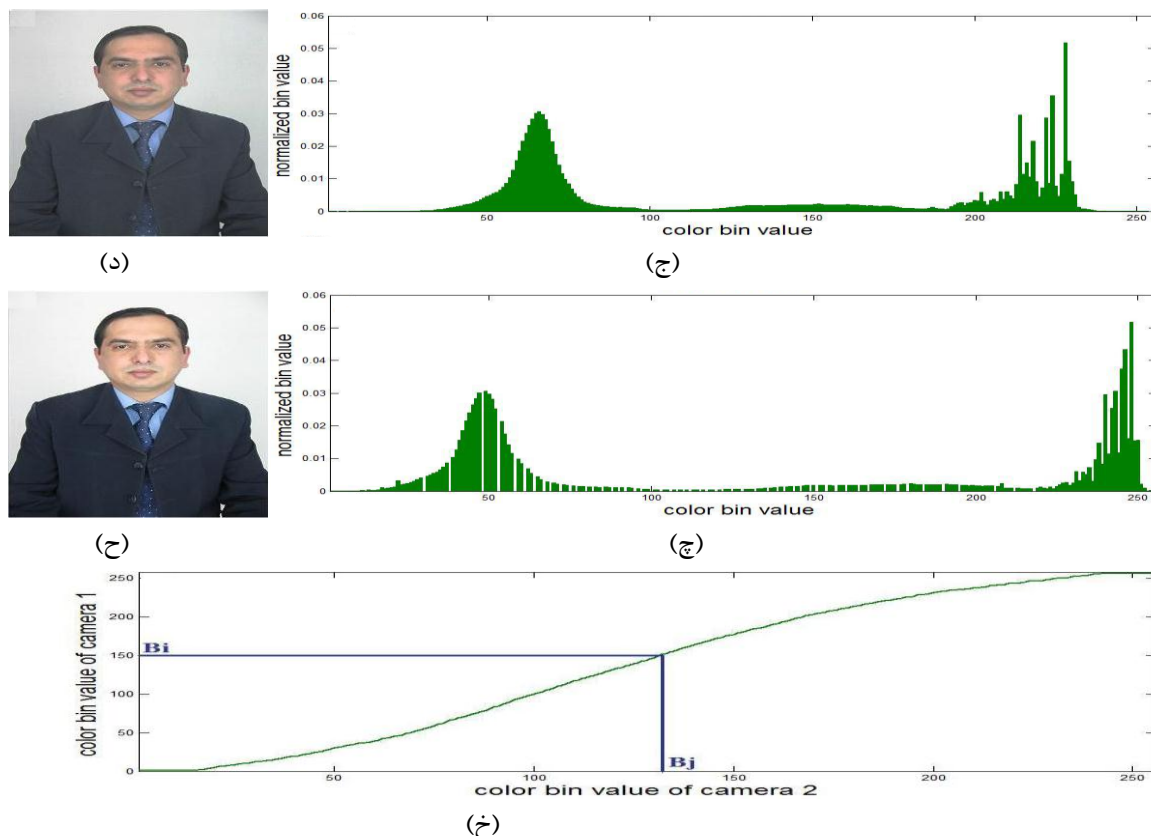
خروجی بدست آمده از این روش تاثیر تغییر شدت روشنایی را تا حد زیادی بر طرف می‌نماید همانطور که در شکل (۴-۱۱) نشان داده شده است این روش در زمانی که از شبکه‌ای از دوربین‌ها استفاده می‌شود و مقدار روشنایی ظاهر افراد زمانی که از دوربینی به دوربین دیگر منتقل می‌شوند بسیار تغییر می‌کند می‌تواند بسیار مفید باشد. مقدار هیستوگرام رنگی که با استفاده از این روش از افراد استخراج شده به عنوان ورودی به تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی که در قسمت بعدی توضیح خواهد داده شده اعمال می‌شود.



(ب)



(الف)



شکل (۴-۱۱) تاثیر شدت روشنایی و تاثیر تابع انتقال روشنایی بر تصاویر. (الف) و (ج) هیستوگرام رنگ سبز تصویر فرد به ترتیب (ب) و (د) در دوربین‌های C_1 و C_2 ، (ب) و (د) تصویر یک فرد در دو دوربین C_1 و C_2 ، (ج) هیستوگرام بعد از اعمال تابع انتقال روشنایی، (ح) تصویر فرد بعد اعمال تابع انتقال روشنایی، (خ) نمایشی از تابع انتقال روشنایی در دو دوربین C_1 و C_2 [۶۴]

۴-۴-۵- تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیر خطی

برای اینکه بتوان انطباق صحیح‌تری بین ویژگی‌های استخراج شده از اشیاء برای شناسایی افراد انجام داد همانطور که در فصل مروری بر کارهای گذشته مطرح شده یکی از روش‌ها، تجزیه و تحلیل مولفه‌های اصلی (PCA) به عنوان یکی از تکنیک‌های با ارزش در این زمینه می‌باشد که این روش، ابزاری برای آنالیز داده‌ها می‌باشد و از آن می‌توان برای کاربردهایی مانند کاهش ابعاد، فشردگی با از دست دادن اطلاعات، استخراج ویژگی و نمایش داده استفاده کرد. این روش به عنوان، تصویر متعامد داده‌ها درون یک فضای خطی با ابعاد کمتر است به قسمی که واریانس بین داده‌های تصویر شده، حداکثر شود [۱۰۳]، که در این روش اگر ما N نمونه با ابعاد D داشته باشیم داده‌ها به فضای با ابعاد M ، $M < D$ به صورتی نگاشت داده می‌شود که در آن واریانس بین داده‌ها جدید حداکثر است. که

اساس کار آن بر اساس مقدار ویژگی می‌باشد که بردارهای ویژگی بر اساس مقدار ویژگی مرتب شده-اند. که می‌توان با توجه به اهمیت ویژگی به انطباق بین افراد بپردازیم.

در این پروژه از روشی به عنوان تجزیه و تحلیل مولفه‌های اصلی قوی فازی که در مرجع [۱۰۴] معرفی شده استفاده نموده‌ایم. در این مرجع از این روش برای دسته‌بندی مجموعه داده‌های پزشکی استفاده شده است که نتایج بدست آمده از ویژگی‌های این روش برای دسته‌بندی رضایت بخش بوده است که به همین منظور ما از آن برای شناسایی مجدد افراد و ردیابی استفاده نموده‌ایم.

الگوریتم تجزیه و تحلیل مولفه‌های اصلی قوی فازی به صورت زیر تعریف می‌شود:

در [۱۰۵] تابع هدف به صورت رابطه (۴-۲۸) پیشنهاد شد که هدف مینیموم کردن $E(U, w)$ با در نظر گرفتن w و u_i می‌باشد.

$$E = \sum_{i=1}^n u_i^{m_1} e(x_i) + \eta \sum_{i=1}^n (1 - u_i)^{m_1} \quad (4-28)$$

اینجا $X = \{x_1, x_2, \dots, x_n\}$ مجموعه داده و $U = \{u_i \mid i = 1, \dots, n\}$ مجموعه عضویت^۱ و η حد آستانه می‌باشد. بازه‌های $u_i \in [0, 1]$ و $m_1 \in [1, \infty)$ می‌باشند u_i مقدار عضویت از x_i که مطلق به مرکز خوشه^۲ می‌باشند و $(1 - u_i)$ مقدار عضویت از x_i که مطلق به نويز (خطا) خوشه و m_1 مقدار فازی بودن^۳ می‌باشد. معیار $e(x_i)$ ، خطا بین x_i و مرکز کلاس می‌باشد. ابتدا مشتق رابطه (۴-۲۹) را نسبت

$$\frac{\partial E}{\partial u_i} = 0 \quad \text{به } u_i \text{ بدست می‌آوریم و با انجام}$$

$$u_i = \frac{1}{1 + \left(\frac{e(x_i)}{\eta}\right)^{1/(m_1-1)}} \quad (4-29)$$

بعد از جایگزینی و ساده سازی رابطه (۴-۳۰) بدست می‌آیند:

^۱ membership set

^۲ Cluster

^۳ Fuzziness

$$E = \sum_{i=1}^n \left(\frac{1}{1 + \left(\frac{e(x_i)}{\eta} \right)^{1/(m_1-1)}} \right)^{m_1-1} e(x_i) \quad (30-4)$$

مشتق نسبت به w خواهیم داشت

$$\frac{\partial E}{\partial w} = \left(\frac{1}{1 + \left(\frac{e(x_i)}{\eta} \right)^{1/(m_1-1)}} \right)^{m_1} \left(\frac{\partial e(x_i)}{\partial w} \right) \quad (31-4)$$

حالا خواهیم داشت

$$\beta(x_i) = \left(\frac{1}{1 + \left(\frac{e(x_i)}{\eta} \right)^{1/(m_1-1)}} \right)^{m_1} \quad (32-4)$$

در مقاله [۱۰۶] PCA غیر خطی را معرفی کرده است که در آن:

$$e_3(x_i) = \|x_i - w^T g(y)\| \quad (33-4)$$

در $y = w^T x_i$ و $g(y)$ تابع غیرخطی می‌باشد. که در این پروژه از تابع $g(y) = \tanh(10y)$ استفاده شده است.

پارمترهای $y = w^T x_i$ ، $u = yw$ و $v = w^T u$ را محاسبه می‌نماییم برای به‌روز رسانی وزن‌ها از رابطه (۳۴-۴) استفاده می‌نماییم.

$$w^{new} = w^{old} + \alpha_T \beta(x_i) (x_i e_3(x_i)^T w^{old} F + e_3(x_i) g(y)) \quad (34-4)$$

در این رابطه α_t ضریب یادگیری می‌باشد و $F = \frac{d}{dy}(g(y))$

الگوریتم پیشنهادی به صورت زیر تعریف می‌شود:

۱. مقدار دهی اولیه $t=1$ ، محدوده تکرار تا T ، ضریب یادگیری $\alpha_0 \in (0,1]$ ، حدآستانه نرم η

مقداردهی مثبت و به صورت تصادفی برای وزن w

۲. تکرار مراحل ۳-۹، تا زمانی که t کمتر از T

۳. محاسبه $\alpha_t = \alpha_0(1-t/T)$ ، و قرار دادن $i=1$ و $\sigma=0$

۴. تکرار ۵-۸، تا زمانی که i کمتر از n

۵. محاسبه $y = w^T x_i$

۶. محاسبه $g(y)$ ، $F = \frac{d}{dy}(g(y))$ ، $e_3(x_i) = \|x_i - w^{old} g(y)\|$ و به روز رسانی وزن

$$w^{new} = w^{old} + \alpha_t \beta(x_i) (x_i e_3(x_i)^T w^{old} F + e_3(x_i) g(y))$$

۷. به روز رسانی مقدار موقت^۱: $\delta = \delta + e_3(x_i)$

۸. $i+1$

۹. محاسبه $\eta = (\delta/n)$ و $t+1$

شکل (۴-۱۲) مراحل کلی الگوریتم تجزیه و تحلیل مولفه‌های اصلی قوی فازی

بعد از این که روش پیشنهادی بر روی مجموعه داده به همراه فرد جدید که وارد زاویه دید دوربین شده پیاده سازی شد. شباهت بین هر فرد با همه افراد مجموعه داده با استفاده از فاصله اقلیدسی که در رابطه (۴-۳۵) نشان داده شده به دست می آید و در صورتی که شباهت این دو فرد از یک حد آستانه (رابطه (۴-۳۶)) کمتر باشد این دو فرد را به عنوان یک شخص واحد معرفی می نماییم در غیر این صورت شخص به عنوان فرد جدید در پایگاه داده قرار می گیرد.

$$dist(o_i, o_j) = \sqrt{\sum_{k=1}^{n*m} (o_{i,k} - o_{j,k})^2} \quad (4-35)$$

$$\min(dist) < Threshold3 \quad (4-36)$$

در رابطه (۴-۳۵) $n*m$ بعد تصویر افراد را نشان می دهد.

¹ Temporary

۴-۵- فضاهای رنگی

پارامترهای زیادی در برقراری ارتباط بین مشاهدات افراد در دوربین‌های مختلف تاثیرگذار می‌باشد که از مهمترین آنها می‌توان به فضاهای رنگی و معیار شباهت اشاره نمود. از آنجای که ما از هیستوگرام رنگ قسمت‌های مختلف بدن افراد برای شناسایی مجدد استفاده نموده‌ایم و از طرفی هیستوگرام فقط به توصیف توزیع کلی رنگ شی می‌پردازد و از جزئیات مکانی و چیدمان رنگ‌ها صرف نظر می‌کند و با توجه به تبدیلاتی که در فضاهای رنگی متفاوت وجود دارد و هریک از مولفه‌ها در این فضاهای رنگی تعریف و تفسیر ویژه‌ای را از خود نشان می‌دهند و می‌تواند در عملکرد ویژگی‌های مبتنی بر رنگ تاثیرگذار باشد، به همین منظور ما در این پایان‌نامه به بررسی تاثیر فضای‌های رنگی در شناسایی-مجدد افراد خواهیم پرداخت. از مهمترین فضاهای رنگی می‌توان به RGB, HSV, YCbCr و ... اشاره نمود.

۴-۶- معیار شباهت

معیارهای زیادی برای مقایسه شباهت‌های بین هیستوگرام رنگ وجود دارد. انتخاب معیار مناسب می‌تواند تاثیر زیادی بر کارایی سیستم‌ردیابی و شناسایی مجدد افراد داشته باشد. از جمله این معیارها می‌توان به فاصله اقلیدسی، ضریب واگرایی kullback-leibler، jensen و Bhattacharyya اشاره کرد. یکی از معیارهای که در انطباق افراد در شبکه‌ای از دوربین‌ها نتایج مناسبی از خود نشان داده معیار Bahhatecheryya می‌باشد [۴۳ و ۱۰۲]. به همین منظور ما از این معیار برای ارزیابی میزان شباهت هیستوگرام افراد استفاده می‌کنیم که در رابطه (۴-۲۱) تعریف شده است.

$$\text{Similarity}(O_{i,k}, O_{j,l}) = 1 - D_B(O_{i,k,c}, O_{j,l,c}) \quad (4-21)$$

در رابطه بالا $O_{i,k,c}$ و $O_{j,l,c}$ نشان دهنده دو مشاهده از افراد در دوربین‌های i و j ، و در فریم‌های k و l می‌باشد در این رابطه c نشان دهنده یکی از فضاهای رنگی قرمز، سبز و آبی می‌باشد. اگر H_i و H_j را هیستوگرام نرمال شده این افراد باشد D_B [۳] به صورت زیر رابطه (۴-۲۲) تعریف می‌شود:

$$D_B(H_i, H_j) = \sqrt{1 - \sum_{v=1}^m \sqrt{H_i(v) \cdot H_j(v)}} \quad (22-4)$$

در رابطه (۲۲-۴) m نشان دهنده تعداد bin در هیستوگرام می‌باشد. رابطه (۲۱-۴) عددی در بازه صفر تا یک را بر می‌گرداند و در صورتی که مقدار یک را برگرداند به معنایی آن می‌باشد که ظاهر آن دو شخص هیچ شباهتی باهم ندارند.

برای شناسایی مجدد افراد، میزان شباهت فرد با افرادی که قبلاً شناسایی شده محاسبه می‌شد، برای این منظور فردی که با توجه به معیار Bahhatecheryya بیشترین شباهت را دارد با استفاده از رابطه (۲۳-۴) بدست می‌آوریم.

$$\arg \max_j \text{similarity}(\text{Similarity}(H_i, H_j)) \quad (23-4)$$

در رابطه (۲۳-۴)، H_i هیستوگرام فردی که تازه وارد زاویه دید دوربین شده و باید شناسایی شود، H_j مربوط به هیستوگرام افراد موجود در پایگاه داده می‌باشد که در این رابطه j به تعداد، افرادی که تا کنون شناسایی شده می‌باشد. در مرحله بعدی میزان شباهت فردی که بیشترین شباهت را با فرد جدید داشته، با در نظر گرفتن یک حد آستانه، با استفاده از رابطه (۲۴-۴) که در [۱۰۲] معرفی شده مقایسه می‌شود.

$$\prod_{ch \in \{R, G, B\}} \gamma^{D(H_i^{ch}, H_k^{ch})} > \text{Threshold} \quad (24-4)$$

در رابطه (۲۴-۴)، D تفاوت هیستوگرام دو فرد می‌باشد و γ عدد ثابت می‌باشد. k ($k \leq j$) فردی می‌باشد که با فرد جدید بیشترین شباهت را دارد. در این رابطه، اگر میزان معیار شباهت برای کانال-های رنگی بیشتر از حد آستانه باشد مبین این مطلب می‌باشد که فرد جدید قبلاً در فریم‌های قبلی مشاهده شده و همان فردی می‌باشد که در رابطه (۲۳-۴) بیشترین انطباق را داشته است (فرد k). اما در صورتی که میزان شباهت یکی از رنگ‌های این دو فرد کمتر از حد آستانه باشد این فرد به عنوان

یک فرد جدید شناخته می‌شود (فرد $j+1$). در این مقاله برای اینکه بتوان مقایسه‌ای بین هیستوگرام قسمت‌های مختلف مثل نیم‌تنه و پایین تنه بدست بیاوریم از رابطه (۴-۲۵) استفاده می‌نماییم.

$$\left(\prod_{ch \in \{R, G, B\}} e^{-\gamma D(H_i^{ch}, H_{k1}^{ch})} \quad \& \quad \prod_{ch \in \{R, G, B\}} e^{-\gamma D(H_i^{ch}, H_{k2}^{ch})} \right) > Threshold \quad (۴-۲۵)$$

در رابطه (۴-۲۵)، $k1$ و $k2$ به ترتیب مبین قسمت‌های، نیم‌تنه و پایین تنه بدن می‌باشد.

شکل (۴-۱۳) مراحل بدست آوردن پارامترهای موقعیت افراد نسبت به دوربین و شناسایی مجدد افراد جدید در شبکه‌ای از دوربین‌ها نشان می‌دهد.

ورودی : تصاویر دوربین‌های C_1 و C_n (تعداد دوربین)

خروجی: ردیابی و شناسایی مجدد افراد

ثابت‌ها: حدآستانه شباهت افراد، حدآستانه‌ی تعداد رکوردها از هر فرد در پایگاه داده

اگر $time < training \text{ time}$ سپس

محاسبه تابع انتقال روشنایی برای دوربین‌های $C_i - C_j$ ($i, j \in \{1, 2, \dots, n\} \delta i \neq j$)

بدست آوردن بیشترین و کمترین ویژگی‌های جدول (۴-۲)

در غیر این صورت

بدست آوردن پارامترهای لازم برای جدول وضعیت با استفاده از رابطه‌های (۴-۷) تا (۴-۱۸) و

موقعیت افراد با استفاده از سیستم فازی

استفاده از تابع انتقال روشنایی برای فرد استخراج شده با توجه به دوربین فعلی و قبلی

انطباق فرد با افراد موجود در حافظه موقت

اگر فرد در حافظه موقت با توجه حدآستانه وجود داشت سپس

به روز رسانی حافظه موقت، پایگاه داده و اضافه شدن در پایگاه داده (در صورتی که تعداد

رکوردهای موجود

از این فرد بیشتر از حد آستانه باشد به جایی قدیمی ترین رکورد ذخیره و در غیر این صورت به پایگاه داده اضافه می‌شود).

در غیر این صورت

انطباق فرد با افراد موجود در پایگاه داده با توجه به موقعیت و ویژگی متناسب با آن

اگر فرد در پایگاه داده وجود داشت

به روز رسانی مجموعه داده‌ها و اضافه شدن در مجموعه داده و حافظه موقت

در غیر این صورت

شناسایی فرد به عنوان شخص

اضافه شدن در مجموعه داده و پایگاه داده

پایان

پایان

پایان

شکل (۴-۱۳) مراحل کلی ردیابی و شناسایی مجدد افراد در شبکه‌ای از دوربین‌ها با دید مجزا

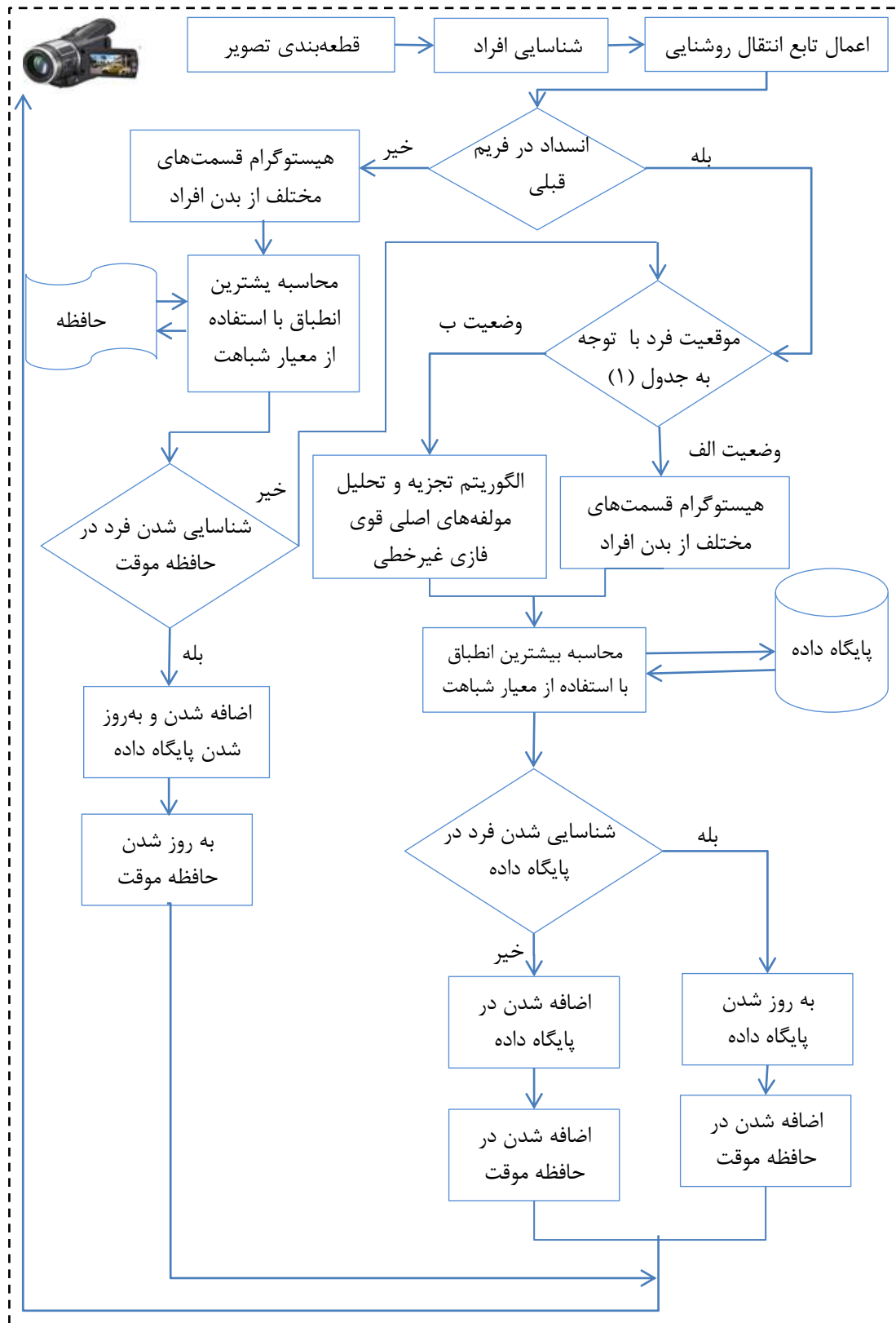
۴-۷- ردیابی در یک دوربین با کمک روش پیشنهادی

سیستمی که در این پایان‌نامه برای شناسایی و ردیابی افراد در یک دوربین معرفی شده نمودار کلی که در شکل (۴-۱۴) نشان داده شده است. در واقع در این قسمت عمل شناسایی را در شبکه‌ای از دوربین‌ها، انجام می‌شود. چون زمانی که در یک دوربین می‌خواهیم عمل شناسایی را انجام دهیم از اطلاعات پایگاه داده‌ای که از دوربین‌های دیگر نیز استخراج شده استفاده می‌نماییم. همانطور که از نمودار کلی مشخص می‌باشد این سیستم قادر به برخورد با مشکل انسداد نیز می‌باشد. که به صورت مختصر به توضیح آن می‌پردازیم.

در این سیستم، در ابتدا افراد با توجه به روش‌های که در فصل قبل معرفی شد از فریم‌ها یا تصاویر دوربین استخراج می‌شوند و سپس با استفاده از تابع انتقال روشنایی تجمعی تاثیر تغییر شدت روشنایی در بین دوربین‌ها کاهش پیدا می‌کند. در گام بعدی بررسی می‌شود که آیا در فریم قبل (حافظه موقت) انسداد رخ داده است یا نه. در صورتی که انسداد رخ ندهد دو امکان ایجاد می‌شود که این فرد در فریم‌های قبلی یا حافظه موقت وجود داشته یا نه. برای این منظور افراد با توجه به ویژگی‌های استخراج شده از هیستوگرام قسمت‌های مختلف از بدن افراد و معیار شباهتی مورد ارزیابی قرار می‌گیرند و در صورتی که فرد جدید با یکی از افراد در حافظه موقت انطباق داشته‌باشد، پایگاه داده و حافظه موقت بروز رسانی می‌شوند در واقع افراد تازه وارد زاویه دید دوربین نشده و فقط در حال ردیابی شدن می‌باشند. اما در صورتی که افراد فریم فعلی با افراد موجود در حافظه موقت انطباق نداشته باشند (این احتمال وجود دارد که حافظه موقت خالی باشد در چند فریم قبلی فردی وجود نداشته باشد). حالتی مشابه به حالتی که انسداد در فریم‌های قبلی رخ داده باشد می‌افتد که در این صورت افراد باید به عنوان افراد جدیدی که در فریم‌های قبلی وجود ندارند شناسایی شوند. اولین کاری که سیستم انجام می‌دهد مشخص کردن موقعیت فرد نسبت به دوربین‌ها می‌باشد که ویژگی‌های استخراج شده برای موقعیت افراد که در پایان‌نامه معرفی شده (طول، عرض، اندازه، فاصله پیکسلی و زاویه) و سیستم‌فازی موقعیت افراد نسبت به دوربین بدست آورده می‌شود. که در این سیستم ما دو حالت کلی از وضعیت‌های افراد نسبت به دوربین‌ها را در نظر گرفتیم که سیستم با توجه به محل قرار گرفتن افراد، از روش پیشنهادی هیستوگرام رنگی قسمت‌های مختلف از بدن افراد یا الگوریتم تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی استفاده می‌نماید. با توجه به این ویژگی‌ها عمل شناسایی با افراد موجود در پایگاه داده انجام می‌شود در صورتی که فرد جدید با افراد موجود در پایگاه‌داده انطباق داشته‌باشد پایگاه‌داده به روزرسانی و فرد جدید به مجموعه‌داده اضافه می‌شود ولی اگر فرد با افراد موجود در پایگاه‌داده انطباق نداشته‌باشد فرد هم به پایگاه‌داده و حافظه

موقت اضافه می‌شود (یعنی در واقع فرد جدیدی وارد سیستم شده که تا کنون در زاویه دید هیچ دوربینی وارد نشده است).

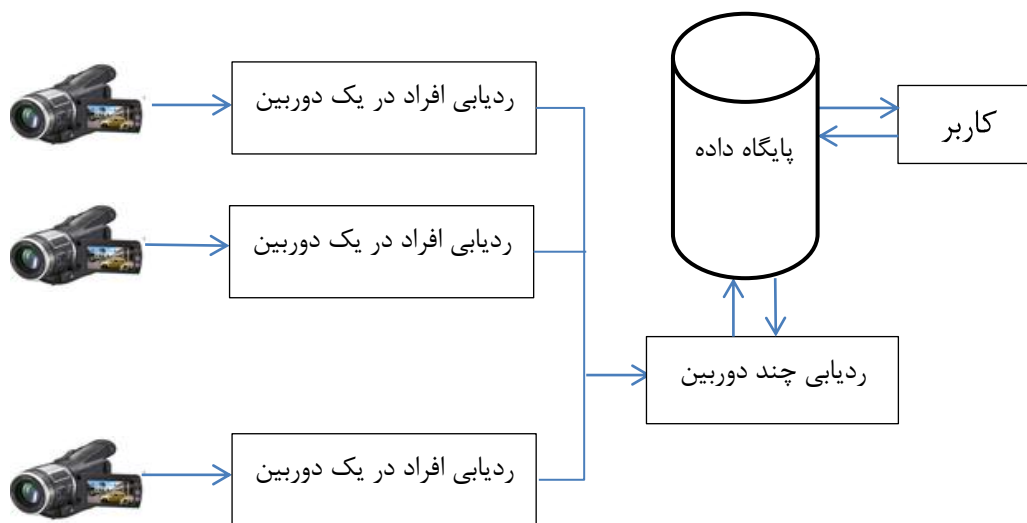
با توجه به مطالب بالا در زمانی که یک فرد جدید وارد زاویه دید یک دوربین می‌شود این فرد در ابتدا با افراد موجود در حافظه موقت و در گام بعدی در صورت وجود نداشتن با افراد موجود در پایگاه داده مقایسه و شناسایی می‌شود یعنی در واقع از چند فریم اول ورود افراد به زاویه دید برای شناسایی استفاده شده ولی این امکان وجود دارد که یک فرد با وجود دیده شدن در زاویه دید دوربین‌های قبلی اما به اشتباه به عنوان فرد جدید معرفی شود و با توجه به الگوریتم، بعد از شناسایی افراد با استفاده از فریم فعلی در فریم‌های بعدی از اطلاعات موجود در حافظه موقت استفاده می‌شود یعنی در واقع امکان تصحیح این خطا وجود ندارد این در حالی می‌باشد سیستم به گونه‌ای طراحی شده که از یک فرد در زمان عبور از یک دوربین نمونه‌های بیشتری قرار گیرد (چند رکورد از یک شخص در پایگاه داده) به همین منظور برای از بین بردن این مشکل، در گام بعدی به صورت موازی در هر چند فریم بعد از این که اطلاعات بیشتری از افراد تازه وارد در پایگاه داده‌ها قرار گرفتن عمل شناسایی مجدد یا در واقع ترکیب افراد احتمالی که به اشتباه شناسایی شده بودند را انجام می‌دهیم برای این منظور از روش ردیابی در شبکه‌ای از دوربین‌ها که در قسمت بعدی توضیح داده شده استفاده می‌نماییم.



شکل (۴-۱۴) شمای کلی از سیستم پیشنهاد به منظور ردیابی و شناسایی مجدد افراد

۴-۸- ردیابی در چند دوربین با زاویه دید مجزا با روش پیشنهادی

برای ردیابی افراد در مجموعه‌ای از دوربین‌ها در این پایان‌نامه از نمودار کلی که در شکل (۴-۱۵) استفاده شده است. زمانی که افراد از یک دوربین به دوربین دیگر منتقل می‌شوند همانطور که در ردیابی و شناسایی افراد در یک دوربین (که این دوربین می‌تواند هر دوربین در سیستم باشد) در قسمت قبل معرفی شد ابتدا با توجه به اطلاعات موجود از افراد در یک یا چند فریم فعلی عمل ردیابی و شناسایی افراد انجام می‌شود و این امکان وجود دارد در هر لحظه از یک دوربین به دوربین دیگر منتقل شد و با توجه به افرادی که در محیط قرار دارند عمل ردیابی صورت گیرد. از آنجایی که افراد فقط با توجه به چند فریم اول که در زاویه دید دوربین‌ها قرار دارند مورد شناسایی و برچسب‌گذاری قرار می‌گیرند و بعد از این مرحله عمل ردیابی صورت می‌گیرد در بعضی از مواقع، به دلیل کم بودن اطلاعات و یا ناقص بودن ظاهر افراد عمل شناسایی به اشتباه رخ دهد و از طرفی در هنگام عبور افراد در صحنه اطلاعات (رکوردهای) بیشتری از هر شخص در پایگاه داده‌ها قرار می‌گیرد به همین منظور، سیستم به صورت موازی هر چند وقت یکبار عمل شناسایی با توجه به این که نمونه‌های بیشتری از افراد در پایگاه داده قرار دارد را انجام می‌دهد در واقع ترکیب شناسه‌های افراد مشابه به هم را انجام می‌دهد. هر نمونه از افرادی را که در پایگاه داده وجود دارند و شباهت زیادی را نسبت به نمونه‌های موجود از افراد دیگر در پایگاه داده دارند را ترکیب نموده و باعث بروز رسانی پایگاه داده نیز می‌شوند. در فصل بعد پیاده‌سازی و نتایج بدست آمده از روش پیشنهادی مورد ارزیابی قرار می‌گیرد.



شکل (۴-۱۵) ساختار سیستم ردیابی انسان برای شبکه‌ای از دوربین با دید مجزا

۹-۴- جمع‌بندی

در این فصل به معرفی مراحل و الگوریتم ردیابی و شناسایی مجدد افراد در شبکه‌ای از دوربین‌ها با دید مجزا پرداخته‌ایم. در ابتدا مراحل کلی در یک سیستم ردیاب مورد بررسی قرار گرفته است و در ادامه روش‌های پیشنهادی برای یک سیستم پرداخته‌ایم. در ابتدا با معرفی جدول وضعیت و ویژگی‌های استخراج از افراد در هر فریم و روش فازی محیط تحت نظارت در هر دوربین به شش وضعیت تقسیم شده است. برای این منظور ویژگی‌های مورد نیاز برای بدست آوردن آن به مانند طول، عرض، ارتفاع، فاصله و جهت افراد نسبت به دوربین‌ها معرفی شده است. با توجه به این، ویژگی‌های جدول قواعد فازی بدست آمده شده، فرایند ایجاد و استنتاج سیستم فازی شرح داده شده است. در ادامه ویژگی هیستوگرام به عنوان ویژگی موثر و سریع برای ردیابی و دوباره‌شناسایی افراد از ظاهر معرفی شده که برای بالابردن قابلیت این روش به تقسیم‌بندی قسمت‌های مختلف بدن انسان با توجه روش پیشنهادی که از فاصله نسبی قسمت‌های مختلف بدن انسان استفاده می‌کند بهره گرفته‌ایم و در انتها روش دیگری به عنوان تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی برای شناسایی مجدد معرفی شده است. هریک از ویژگی‌ها با توجه به موقعیت قرار گرفتن نسبت به دوربین مورد استفاده

قرار گرفته شده است. در فصل بعد روش پیشنهاد شده در این فصل با جزئیات پیاده‌سازی تشریح می‌شود.

فصل پنجم

پیاده سازی و ارزیابی نتایج

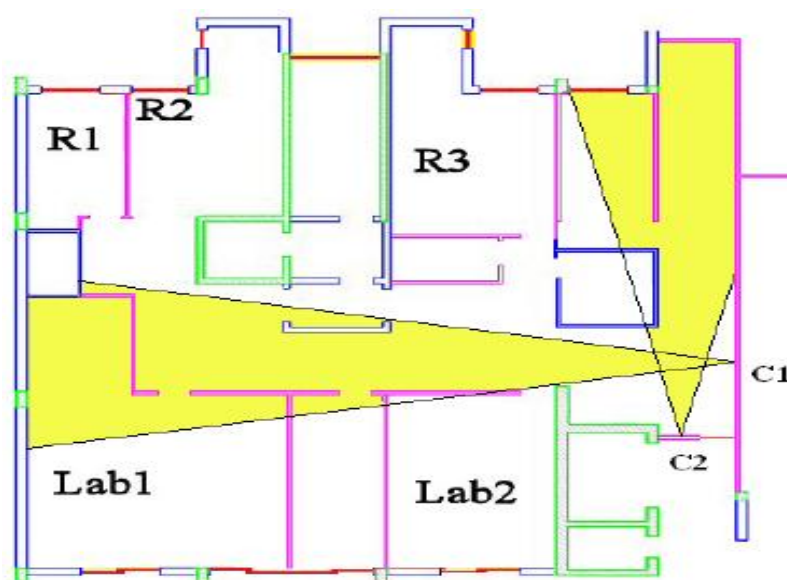
۵-۱- مقدمه

در این فصل به بررسی پیاده‌سازی‌های انجام شده از مباحث و مفاهیم ارائه شده تا کنون و نتایج حاصل از آنها پرداخته خواهد شد. تمرکز این فصل بیشتر بر روی سیستم معرفی شده در این پایان-نامه و مراحل طراحی و نتایج حاصل از آن می‌باشد. بدین منظور در ابتدا پایگاه داده استفاده شده و پیچیدگی‌های موجود در آن معرفی می‌شود. در گام بعدی مراحل لازم برای استخراج ویژگی‌های معرفی شده هم به منظور سیستم‌سازی جهت تقسیم‌بندی محیط تحت مراقبت و ویژگی‌های لازم برای ردیابی به صورت مرحله به مرحله انجام می‌شود و در نهایت نتایج حاصل از سیستم پیشنهادی بر روی پایگاه داده مورد تجزیه و تحلیل قرار گرفته خواهد شد.

همانطور که قبلاً ذکر شد، عمل مشاهده یک شی در زاویه دید یک دوربین و شناسایی مجدد همان شی در زاویه دید همان دوربین یا دوربین دیگر را شناسایی مجدد شی گویند. یک شی ممکن است در زاویه دید دوربین، یک یا چندبار وارد شود. اگر بتوان کار شناسایی مجدد یک شی و اطلاعات مربوط به زمان و تعداد دفعات ورود یا خروج آن شی از زاویه دید یک دوربین را بدست آورد، کمک بسیار مهمی برای شناسای و آنالیز رفتار اشیاء انجام می‌گیرد. برای ردیابی اشیاء در محیط‌هایی که از شبکه-ای از دوربین‌ها با دید مجزا استفاده می‌شود ویژگی حرکت نمی‌تواند برای شناسایی مجدد شی مفید باشد. زیرا در این نوع از محیط‌ها نقاط ورود و خروج زیادی وجود دارند این در حالی است که ویژگی حرکت به مکان شی بستگی دارد. این امکان وجود دارد که اشیاء از یک مکان خارج شوند و دوباره در همان صحنه، ولی این بار از مکان دیگر، وارد شوند. شکل (۱-۱) حالت‌های مختلف از ورود و خروج افراد را در این نوع از سیستم‌ها نشان می‌دهد که با توجه به متغیر بودن زمان و مکان اشیاء در هنگام ورود و خروج، می‌توان نتیجه گرفت که روش‌های مبتنی بر زمان-مکان نمی‌توانند مناسب باشند.

به طور مشابه خواص هندسی نمی‌توانند نتایج خوبی برای شناسایی دوباره اشیاء برای اشیاء غیر صلب به دلیل، کیفیت پایین تصاویر و تغییر اندازه بسیار زیاد (دور و نزدیک شدن شی نسبت به دوربین‌های

متفاوت)، از خود نشان دهند [۶۴]. ولی ویژگی‌های مبتنی بر مدل ظاهری افراد دارای استحکام (قدرت بالا) در تشخیص جسم صلب و غیرصلب می‌باشد به همین منظور در این پایان از این نوع ویژگی برای شناسایی مجدد افراد استفاده شده است.



شکل (۵-۱) نمای کلی از محیط تحت نظارت توسط دو دوربین با زاویه دید مجزا [۱۰۲]

۵-۲- پایگاه داده

در این پایان‌نامه از مجموعه داده‌ی معرفی شده در مقاله [۱۰۲] استفاده شده است. شکل (۵-۱) نقشه‌ای از محیط تحت مراقبت این مجموعه داده را نشان می‌دهد که در آن دو دوربین در قسمت متفاوت یک اداره نصب شده و زاویه دید این دوربین‌ها هیچ هم‌پوشانی باهم ندارد. تعداد انسان‌هایی که در این مجموعه وجود دارد ۱۵ نفر می‌باشد. که ۱۲ مرد و ۳ زن می‌باشند. همانطور که در شکل (۱-۱) نشان داده شد و در فصل اول گفته شد در این پایان‌نامه تمام حالت‌های ممکن در عبور افراد در بین دو دوربین در نظر گرفته شده است. جدول‌های (۵-۱) تا (۵-۴) نمونه‌هایی از افراد موجود در این مجموعه داده را در حالت‌های متفاوت نشان می‌دهد. بعضی از این افراد فقط در یک دوربین ظاهر می‌شوند و چند بار وارد زاویه یک دوربین می‌شوند و بعضی دیگر از افراد در زاویه دید هر دو دوربین ورود و خروج دارند.

جدول (۵-۱) افرادی را نشان می‌دهد که در زاویه دید هر دو دوربین مشاهده می‌شوند و این افراد در حالت‌های متفاوت از زاویه دید دوربین‌ها وارد و خارج می‌شوند مثلاً اینکه یک فرد چند بار وارد زاویه دید یک دوربین می‌شود و بعد از چند مرحله حضور در یک دوربین، همان فرد در زاویه دید دوربین دیگر وارد می‌شود این افراد تمام حالت‌های ممکن در رفت و آمد افراد در بین دو دوربین را دارند.

جدول (۵-۱) افراد موجود در مجموعه داده که در زاویه دید دو دوربین در حال رفت و آمد هستند

دوربین ۱	دوربین ۲	دوربین ۱	دوربین ۲	دوربین ۱	دوربین ۲	دوربین ۱	دوربین ۲	دوربین ۱	دوربین ۲
				فرد ۲				فرد ۳	
				فرد ۴				فرد ۵	
									

جداول (۲-۵) و (۳-۵) حالت‌های دیگر در عبور افراد از زاویه دید دوربین‌ها را نشان می‌دهد این جداول افرادی را نشان می‌دهند که در زاویه دید فقط یکی از دوربین‌ها دیده شده‌اند افراد در زاویه دید هر دوربین در حال رفت و آمد می‌باشند در واقع هر فرد در زاویه دید یک دوربین از هر مکان از محیط وارد شده و از مکان دلخواه خارج می‌شود و در مکان دیگر در همان محیط می‌تواند دوباره وارد دوربین می‌شود. تمام حالت‌های متفاوت در هنگام ورود و خروج از زاویه دید یک دوربین را این افراد دارند. جدول (۲-۵) افراد موجود در دوربین شماره یک و جدول (۳-۵) افراد موجود در دوربین شماره دو را نشان می‌دهد.

جدول (۲-۵) افراد موجود در مجموعه داده که در زاویه دید دوربین شماره یک در حال رفت و آمد هستند

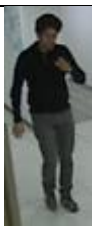
دوربین ۱	دوربین ۱		دوربین ۱	دوربین ۱	
		فرد ۷			فرد ۶

جدول (۳-۵) افراد موجود در مجموعه داده که در زاویه دید دوربین شماره دو در حال رفت و آمد هستند

دوربین ۲	دوربین ۲		دوربین ۲	دوربین ۲		دوربین ۲	دوربین ۲	
		فرد ۱۰			فرد ۹			فرد ۸

جدول (۴-۵) افراد موجود در دو دوربین را نشان می‌دهد که این افراد فقط یک بار در زاویه دید فقط یک دوربین مشاهده می‌شوند. حالت‌های ورود و خروج افراد در یک دوربین فقط برای یکبار متفاوت می‌باشد.

جدول (۴-۵) افراد موجود در مجموعه داده که در زاویه دید فقط یکی از دو دوربین در حال عبور می‌باشند

دوربین ۱	دوربین ۱	دوربین ۲	دوربین ۲	دوربین ۲	دوربین ۲	دوربین ۲	دوربین ۲
فرد ۱۱		فرد ۱۲		فرد ۱۳		فرد ۱۴	
						فرد ۱۵	

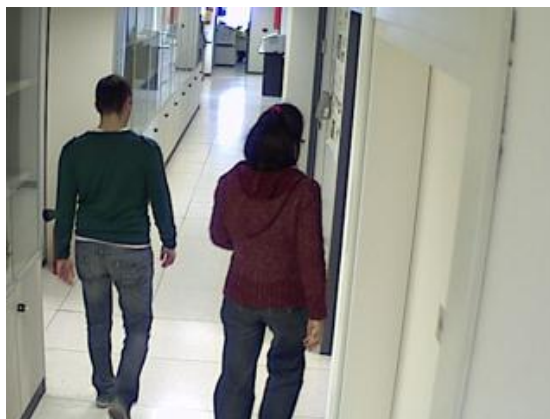
شکل (۲-۵) افراد موجود در مجموعه داده پس از حذف پس‌زمینه را نشان می‌دهد.



شکل (۲-۵) افراد موجود در مجموعه داده - ۱۱ مرد و ۳ زن

اندازه تصاویر در این پایگاه داده 840×640 پیکسل بوده که برای بالا بردن سرعت ما از تصاویر با ابعاد 320×240 استفاده کرده‌ایم. شکل (۳-۵) دو نمونه از محیط تحت مراقبت این مجموعه داده را نشان

می‌دهد



شکل (۳-۵) دو نمونه از تصاویر مجموعه داده در دو دوربین متفاوت

افراد در این مجموعه داده، با توجه به نسبت فاصله‌ی افراد تا دوربین اندازه‌های متفاوتی دارند که در شکل (۴-۵) دو نمونه از این افراد با اندازه‌های متفاوت در دو دوربین را نشان می‌دهد.



شکل (۴-۵) نمونه‌ای از دو فرد با اندازه‌های متفاوت در دو دوربین

۵-۳- بدست آوردن پارامترهای سیستم ردیاب در مرحله راه‌اندازی

برای بدست آوردن پارامترهای مورد نیاز برای جدول وضعیت (قواعد فازی) و همین طور برای بدست آوردن تابع انتقال روشنایی در مرحله آموزش یا راه‌اندازی یک فرد از زاویه دید دو دوربین در محیط-های متفاوت عبور می‌کنند و مقدار بیشترین و کمترین، ویژگی‌های مطرح شده در فصل قبل برای جدول وضعیت (ارتفاع، عرض، اندازه و فاصله پیکسلی) را در هر دوربین محاسبه می‌نماییم. در قسمت دیگر مقدار تابع انتقال روشنایی را در هنگام عبور از دوربین یک به دوربین دو و بالعکس بدست می‌آوریم. با توجه به آزمایشات انجام شده مقدار بیشینه و کمینه ویژگی‌های مطرح شده در جدول (۵-۵)

نشان داده شده است. لازم به ذکر می‌باشد که از تصاویر مربوط به فرد ۵ در هنگام عبور از زاویه دید دو دوربین استفاده شده است و این فرد در مجموعه آزمون یا تست از زاویه دید دوربین‌ها عبور نمی‌کند.

جدول (۵-۵) مقادیر کمینه و بیشینه ویژگی‌های جدول وضعیت

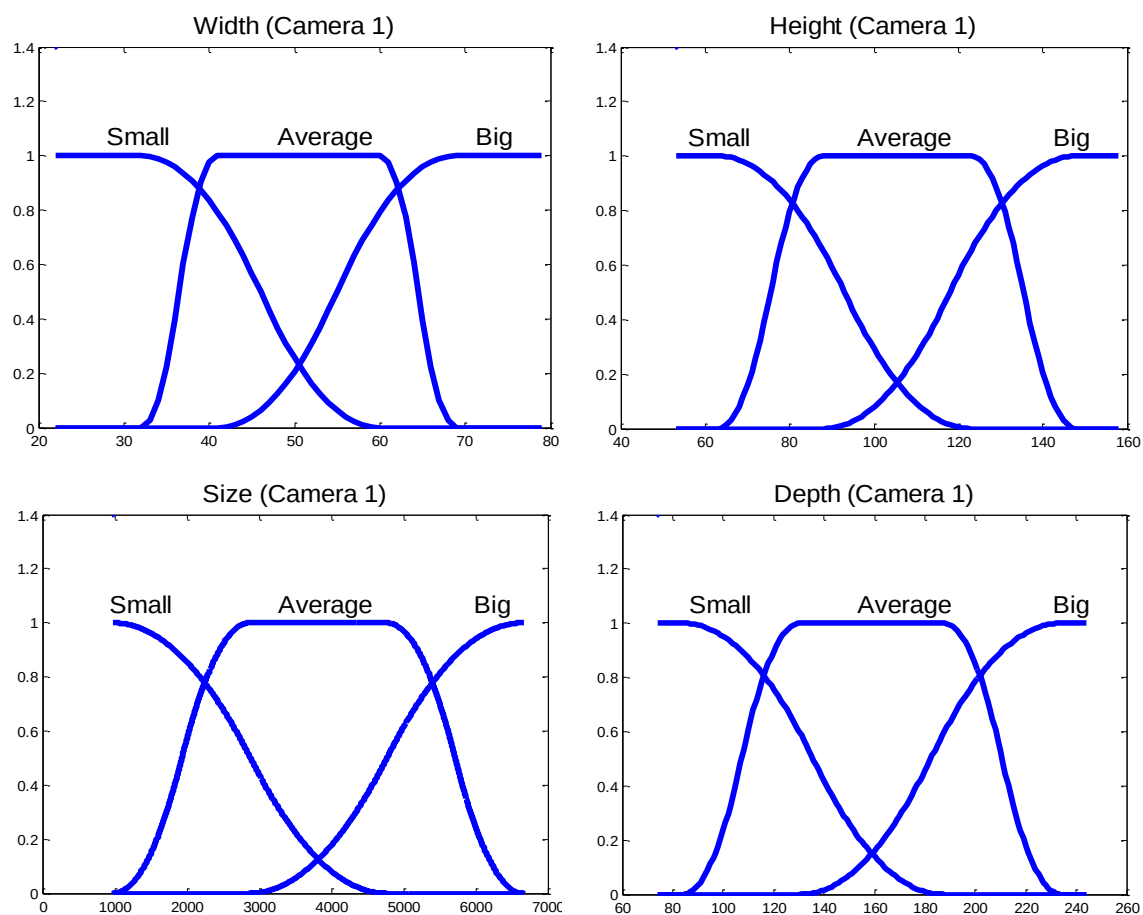
ارتفاع	عرض	اندازه	فاصله	
۵۳	۲۲	۹۷۸	۷۴	کمینه دوربین ۱
۱۵۸	۷۹	۶۶۷۲	۲۴۴	بیشینه دوربین ۱
۵۱	۲۳	۱۰۲۲	۵۴	کمینه دوربین ۲
۲۰۷	۱۱۷	۱۴۱۱۳	۲۵۳	بیشینه دوربین ۲

با توجه به مقادیر بدست آمده از جدول (۵-۵) می‌توان جدول وضعیت و ورودی‌های سیستم فازی را محاسبه نمود که مقادیر بدست آمده برای پارمترهای ورودی فازی با استفاده از رابطه‌های (۴-۷) تا (۴-۱۸) به صورت زیر به ترتیب برای دوربین‌های ۱ و ۲ می‌باشد.

جدول (۶-۵) مقادیر بدست آمده برای جدول وضعیت در دو دوربین

	دوربین ۱				دوربین ۲			
	ارتفاع	عرض	اندازه	فاصله	ارتفاع	عرض	اندازه	فاصله
Small	۸۳<	۴۱<	۲۸۷۶<	۱۳۰<	۱۰۳<	۵۴<	۵۳۸۵<	۱۲۰<
Average	۸۳>	۴۱>	۲۸۷۶>	۱۳۰>	۱۰۳>	۵۴>	۵۳۸۵>	۱۲۰>
	&	&	&	&	&	&	&	&
	<۱۲۳	<۶۰	<۴۷۷۴	<۱۸۷	<۱۵۵	<۸۵	<۹۷۴۹	<۱۸۶
Big	>۱۲۳	>۶۰	>۴۷۷۴	>۱۸۷	>۱۵۵	>۸۵	>۹۷۴۹	>۱۸۶

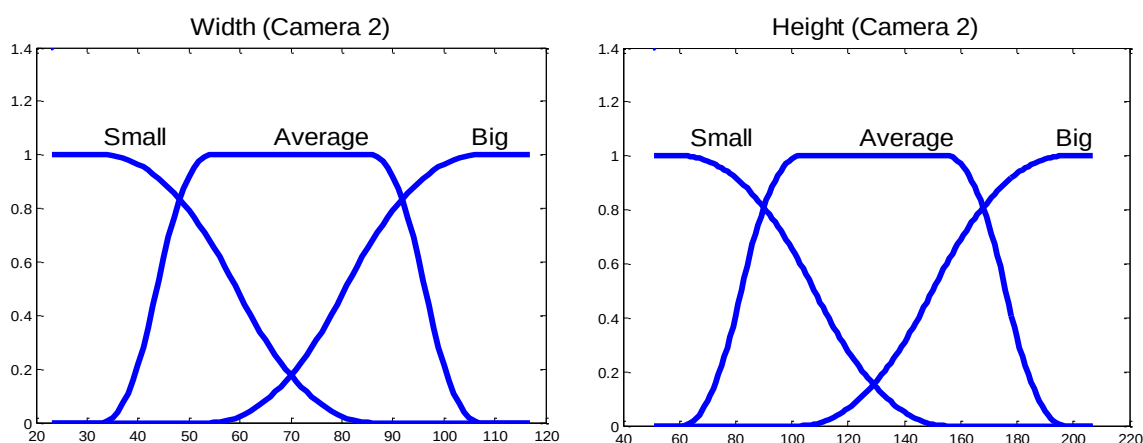
تابع‌های فازی برای هر ویژگی با استفاده از مقادیر جدول (۶-۵) برای دوربین شماره یک به صورت شکل (۵-۵) می‌باشد.

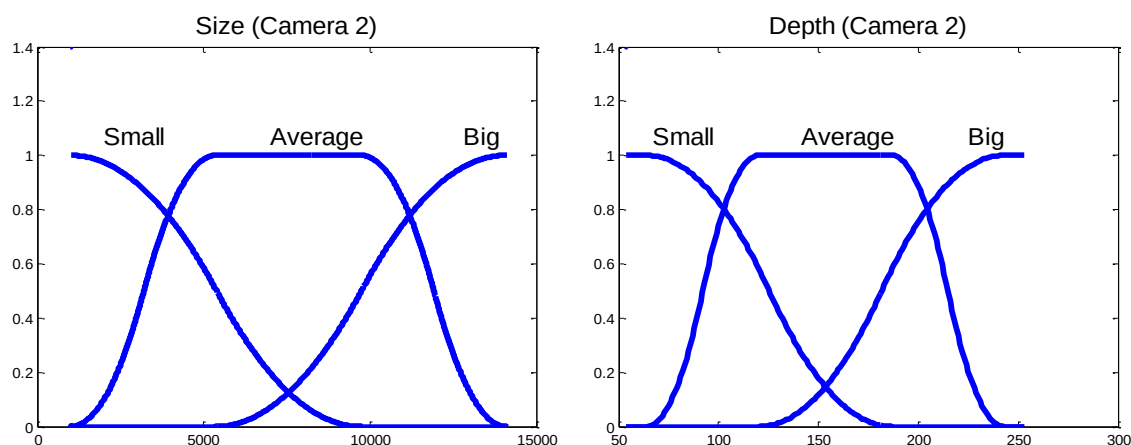


شکل (۵-۵) توابع فازی با استفاده از مقادیر جدول (۵-۶) برای دوربین شماره یک

تابع‌های فازی برای هر ویژگی با استفاده از مقادیر جدول (۵-۶) برای دوربین شماره دو به صورت

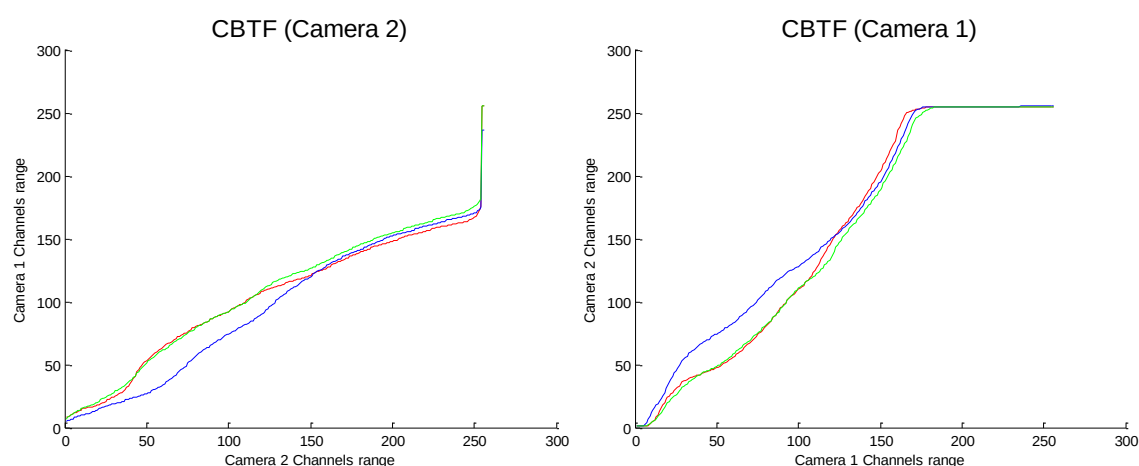
شکل (۶-۵) می‌باشد.





شکل (۶-۵) توابع فازی با استفاده از مقادیر جدول (۶-۵) برای دوربین شماره دو

تابع انتقال روشنایی تجمعی معکوس بدست آمده از فرد شماره پنج در هنگام عبور از دوربین یک به دو و بالعکس به صورت شکل (۷-۵) می‌باشد.



شکل (۷-۵) تابع انتقال روشنایی بدست آمده برای دو دوربین در مرحله آموزش. الف) تابع انتقال روشنایی برای دوربین شماره یک، ب) تابع انتقال روشنایی برای دوربین شماره دو.

۴-۵- نتایج بدست آمده از وضعیت افراد نسبت به دوربین‌ها

با توجه به پارامترهای بدست آمده از آزمایشات قبل می‌توان وضعیت افراد نسبت به دوربین‌ها را بدست آورد به همین منظور در شکل (۸-۵) نمونه‌های از وضعیت افراد در دو محیط جدا نسبت به دوربین‌ها را نشان می‌دهیم. همانطور که از جدول وضعیت مشخص می‌باشد وضعیت افراد نسبت به هر دوربین به شش قسمت تقسیم می‌شود که سه قسمت زمانی که افراد روبه دوربین و سه قسمت دیگر

عکس آن در زمانی که پست به دوربین می‌باشد.



الف

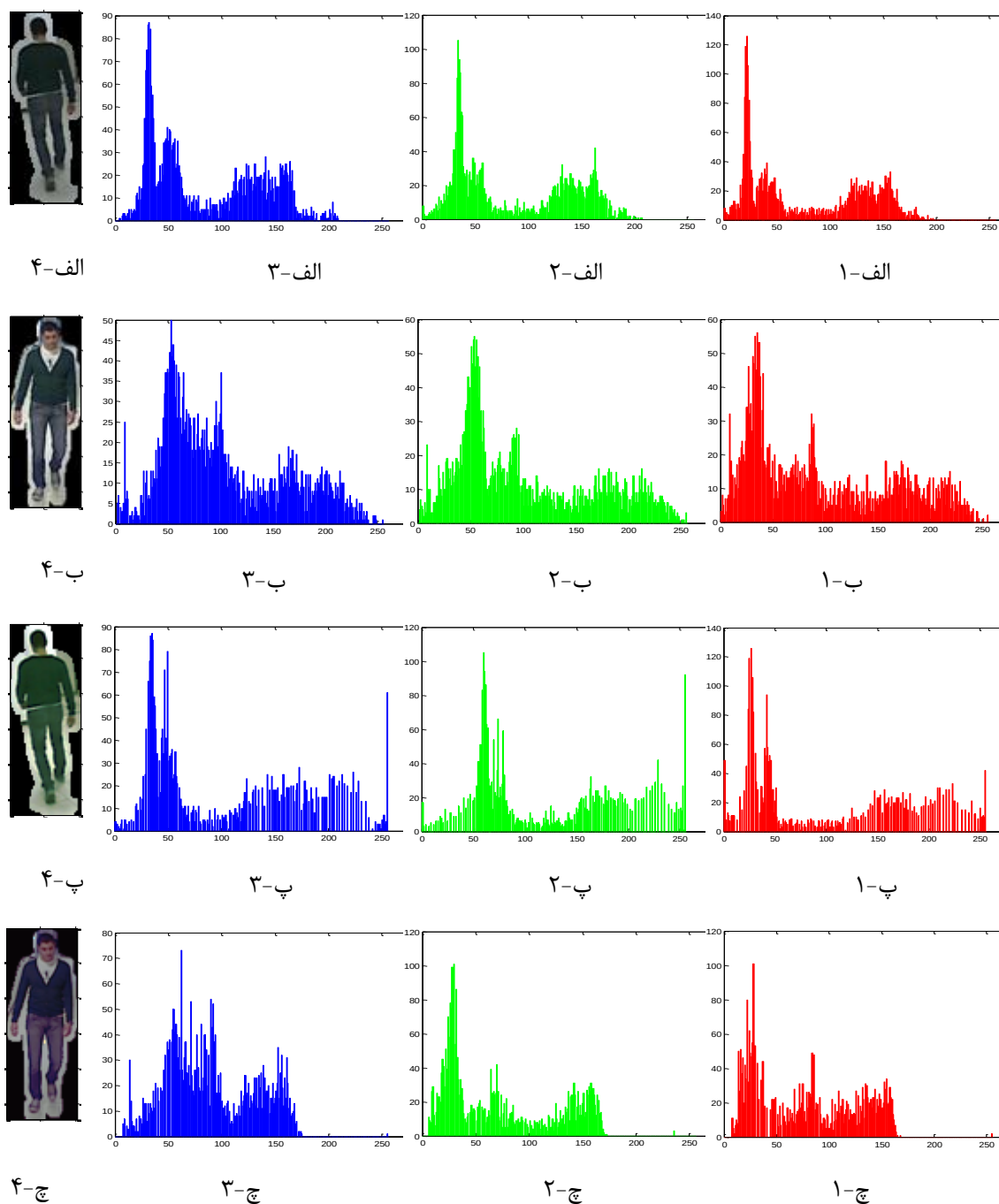


شکل (۵-۸) نتایج بدست آمده از روش فازی به منظور تعیین موقعیت افراد نسبت به دوربین‌ها. الف) موقعیت افراد در

دوربین ۱ و ب) موقعیت افراد در دوربین ۲

۵-۵- نتایج بدست آمده از تابع انتقال روشنایی تجمعی

در این پایان‌نامه از ویژگی هیستوگرام برای دوباره‌شنایی افراد در وضعیت‌های ۱، ۳، ۴ و ۶ استفاده شده است که یکی از مشکلات استفاده از هیستوگرام افراد برای شناسایی مجدد، تغییر شدت روشنایی می‌باشد و از طرفی که در این پایان‌نامه از دوربین‌های با دید مجزا استفاده شده و این امکان وجود دارد که شدت روشنایی محیط در قسمت‌های مختلف بسیار با هم متفاوت باشد به همین منظور ما از یک تابع انتقال روشنایی که در فصل قبل در مورد آن به صورت کامل توضیح داده شده به عنوان تابع انتقال روشنایی تجمعی معکوس استفاده نموده‌ایم. شکل (۵-۹) نمونه‌ای از نتایج بدست آمده از این تابع انتقال روشنایی را برای یک فرد در دو دوربین متفاوت را نشان می‌دهد. در این شکل هیستوگرام کانال‌های مختلف رنگی به صورت جدا برای این فرد در دوربین‌های متفاوت نشان داده شده است. همانطور که می‌توان در هیستوگرام‌ها مشاهده نمود تصویر این فرد از لحاظ شدت روشنایی باهم متفاوت می‌باشند در واقع بازه‌ی تغییرات رنگ در دو تصویر از یک فرد متفاوت می‌باشد که این امر به دلیل تغییر شرایط نوری محیط می‌باشد. به همین منظور ما از تابع انتقال روشنایی برای کاهش این تغییرات استفاده نموده‌ایم. شکل (پ-۴) تصویر فرد (الف-۴) می‌باشد که بعد از اعمال تابع انتقال مربوط به دوربین ۱ بر روی رنگ این تصویر را نشان می‌دهند همانطور که در هیستوگرام این تصویر نشان داده شده بازه رنگی این فرد با بازه رنگی شکل (ب-۴) نزدیک تر می‌باشد. و از طرف دیگر شکل (چ-۴) تصویر فرد (ب-۴) می‌باشد که بر روی تصویر این فرد نیز به صورت مشابه از تابع انتقال روشنایی مربوط به دوربین ۲ استفاده شده است. همانطور که در هیستوگرام این تصویر نشان داده شده بازه رنگی این فرد با بازه رنگی شکل (الف-۴) نزدیک تر می‌باشد.



شکل (۹-۵) تاثیر تابع انتقال روشنایی در هنگام عبور افراد در بین دوربین‌ها. (الف-۴ و ب-۴) تصویر فرد استخراج شده از دوربین‌ها به ترتیب ۱ و ۲. (الف-۱)، (الف-۲)، (الف-۳)، (ب-۱)، (ب-۲) و (ب-۳) هیستوگرام کانال‌های رنگی قرمز، سبز و آبی برای فرد استخراج شده از دوربین یک و دو. (پ-۴) تصویر فرد در دوربین یک بعد از اعمال تابع انتقال روشنایی مربوط به دوربین یک. (چ-۴) تصویر فرد در دوربین دو بعد از اعمال تابع انتقال روشنایی مربوط به دوربین دوم. (پ-۱)، (پ-۲)، (پ-۳)، (چ-۱)، (چ-۲) و (چ-۳) هیستوگرام کانال‌های رنگی قرمز، سبز و آبی برای فرد بعد از اعمال تابع انتقال روشنایی.

۵-۶- محاسبه تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی

در این پایان‌نامه از روش تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی بر روی هیستوگرام رنگ افراد برای دوباره‌شنایی در وضعیت‌های ۲، ۵ نسبت به محیط استفاده شده است. برای استفاده از این روش برای استخراج بهترین ویژگی‌ها از افراد موجود در پایگاه‌داده‌ها و فرد جدید استفاده شده است که بعد از بدست آوردن ویژگی‌ها با استفاده از الگوریتم شکل (۴-۱۳) به شناسایی و دوباره‌شنایی افراد می‌پردازیم البته در شایان به ذکر است که ورودی این الگوریتم میانگین هر سه کانال رنگی می‌باشد.

۵-۷- معیارهای ارزیابی

دو معیار دقت^۱ و یادآوری^۲ به عنوان معروف‌ترین معیار برای شناسایی مجدد افراد در کارهای که در سال‌های اخیر انجام شده، مورد استفاده می‌باشد. به همین منظور در این پایان‌نامه برای ارزیابی بهتر روش پیشنهادی از این دو معیار استفاده نموده‌ایم. رابطه‌های استفاده شده به صورت زیر می‌باشد.

$$PR = \frac{TP}{TP + FP} \quad (۱-۵)$$

$$RE = \frac{TP}{TP + FN} \quad (۲-۵)$$

True positive (TP): اگر یک برچسب قدیمی به درستی به یک شی نسبت داده شود یا برچسب جدید به یک شی جدید داده شود.

False positive (FP): اگر یک برچسب قدیمی به به یک فرد به اشتباه نسبت داده شود.

False negative (FN): اگر برچسب جدید برای یک شی که قبلاً شناسایی شده و دارای برچسب بوده

یکی دیگر از معیارهای که به عنوان میانگین دو معیار بالا مطرح شده معیار F1، رابطه (۵-۳) می‌باشد.

^۱ Percision

^۲ Recall

$$F1 = \frac{2 \times PE \times RE}{PE + RE} \quad (3-5)$$

یکی دیگر از معیارهای که در بعضی از کاربردها مورد ارزیابی قرار گرفته معیار Accuracy می‌باشد که نسبت تعداد افرادی که به درستی شناسایی شده به کل افراد را محاسبه می‌نماید. البته شایان به ذکر است که در این پایان نامه برای سه معیار اول معرفی شده در صورتی که برای یک فرد به اشتباه دو برچسب در پایگاه داده در نظر گرفته شده باشد و فرد دوباره وارد زاویه دید دوربین‌ها شود اگر سیستم برچسب فرد جدید را یکی از این دو برچسب در نظر گیرد سیستم به درستی عمل شناسایی مجدد را انجام داده است. ولی برای معیار Accuracy اولین برچسب که برای هر فرد منتسب شده مهم می‌باشد. مثلاً اگر برای فرد a دو برچسب ۱ و ۲ نسبت داده شده باشد یعنی برای این شخص دو رکورد متفاوت قرار داده شده که در اینجا یک اشتباه رخ داده است ولی اگر همین شخص a دوباره وارد زاویه دید یکی از دوربین‌ها شود و هر یک از برچسب‌های ۱ یا ۲ به آن نسبت داده شود درست عمل شناسایی مجدد رخ داده است. ولی در معیار Accuracy فقط برچسب ۱ تا پایان شناسایی مجدد باید برای این فرد در نظر گرفته شود.

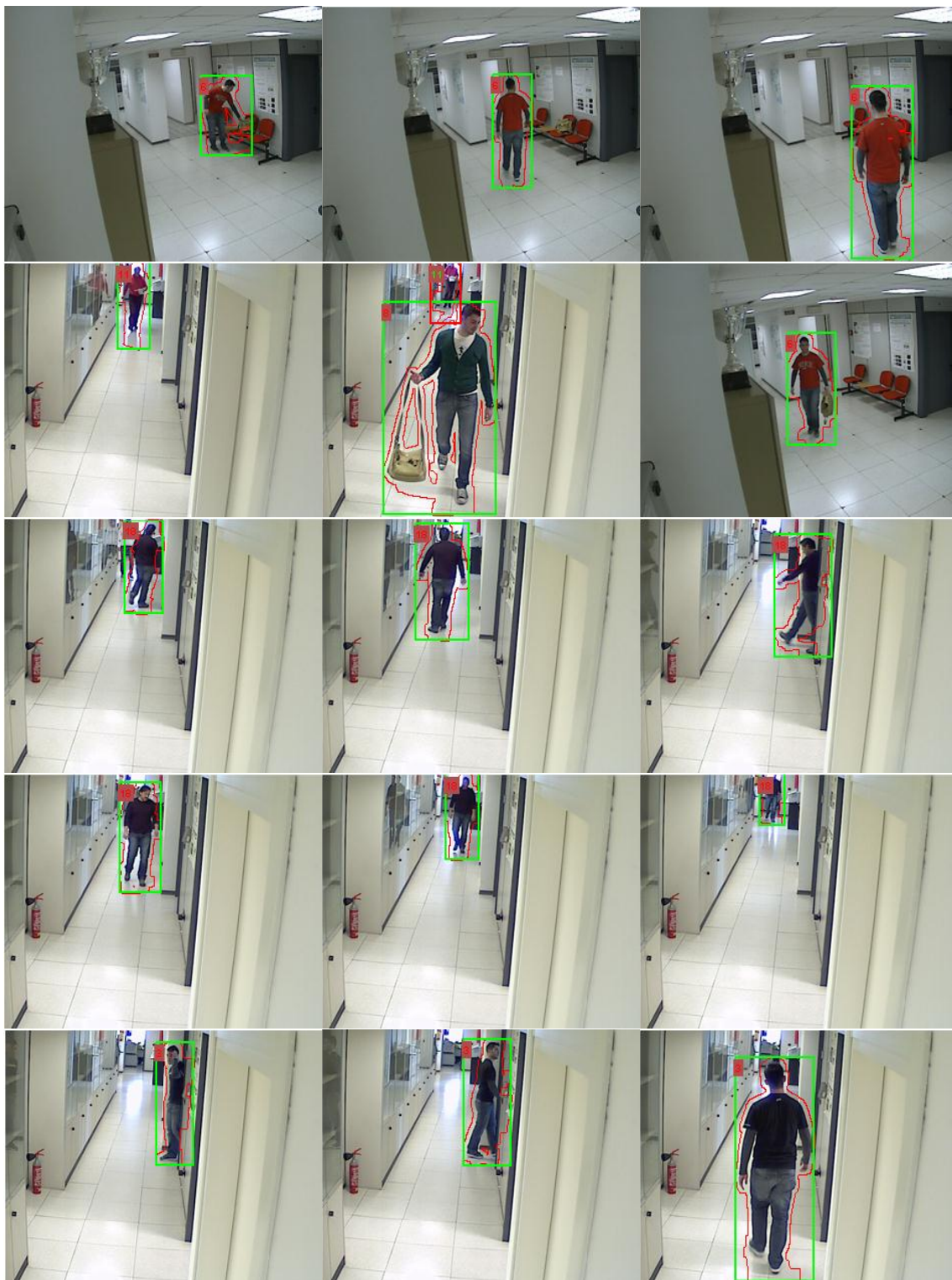
۵-۸- نتایج بدست آمده برای ردیابی انسان در شبکه‌ای از دوربین‌ها

جهت ارزیابی الگوریتم‌های ارائه شده برای ردیابی افراد در شبکه‌ای از دوربین‌ها با دید مجزا، در مجموع از ۲۶۴۲ توالی مختلف از مجموعه داده معرفی شده استفاده نموده‌ایم همانطور که معرفی شده ۱۵ فرد مختلف در حال رفت و آمد می‌باشند از این مجموعه داده ۲۰۸ توالی برای مرحله آموزش در نظر گرفته شده که شخص شماره ۵ در این مرحله از زاویه دید دو دوربین عبور می‌کند (۱۰۴ توالی در دوربین شماره ۱ و ۱۰۴ توالی در دوربین شماره دو). از ۲۴۳۴ توالی دیگر برای مرحله آزمون استفاده شده است که ۱۴ فرد (بدون فرد استفاده شده در مرحله آموزش) در این مرحله از زاویه‌های دید دوربین‌های مختلف عبور کرده و محل ورود و خروج این افراد در مجموع متفاوت می‌باشد. برای ردیابی افراد در این سیستم از روش‌های که در این پایان‌نامه معرفی شده استفاده نموده‌ایم برای این

منظور با توجه به موقعیت افراد نسبت به دوربین از ویژگی‌های مربوط به هر موقعیت استفاده شده است در ابتدا با توجه به این که افراد در زاویه دید کدام دوربین قرار دارند از تابع انتقال روشنایی برای از بین بردن تغییرات شدت روشنایی در هنگام عبور افراد از بین دوربین‌ها استفاده شده و در گام بعدی با استفاده از روش فازی، موقعیت هر فرد نسبت به دوربین‌ها بدست آمده و برای شناسایی افراد با توجه موقعیت آنها، که اگر در وضعیت‌های یک، سه، چهار و شش از روش هیستوگرام قسمت‌های مختلف و در وضعیت‌های دو و پنج از روش تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیرخطی استفاده شده است. شکل (۵-۱۰) نتیجه‌ی به دست آمده برای ردیابی افراد در این سیستم را به صورت تصویری نشان می‌دهد. نتایج نشان دهنده عملکرد قابل قبول سیستم ردیابی در شبکه‌ای از دوربین‌ها می‌باشد.









شکل (۵-۱۰) نمونه‌هایی از نتایج شناسایی و ردیابی افراد با روش‌های پیشنهادی در شبکه‌ای از دوربین‌ها

۵-۹- نتایج شناسایی مجدد افراد

همانطور که در فصل‌های قبل توضیح داده شد یکی از پارامترهای موثر برای بهبود عملکرد سیستم ردیاب، فضای رنگی مورد استفاده می‌باشد به همین منظور در این پایان‌نامه آزمایشات با استفاده از روش‌های پیشنهادی در فضاهای رنگی متفاوت مورد ارزیابی قرار گرفته است. نتایج بدست آمده از آزمایشات نشان دهنده بهبود عملکرد سیستم پیشنهادی می‌باشد و از طرفی فضای رنگی YCbCr بهترین عملکرد را با ۹۱٪ معیار F1 و ۶۵٪ معیار Accuracy داشته است. البته باید به این نکته نیز اشاره کرد که آزمایشات مقاله [۱۹] تنها برای چهار فرد در این مجموعه داده انجام شده ولی ما در این پایان‌نامه آزمایشات را برای شناسایی مجدد چهارده نفر انجام داده‌ایم.

جدول (۵-۷) نتایج بدست آمده با استفاده از روش پیشنهادی برای شناسایی دوباره افراد در شبکه‌ای از دوربین‌ها با دید مجزا در محیط بسته

Accuracy	F1	RE	PR	روش‌ها
۵۸٪	۷۶٪	۷٪	۸۳٪	روش پیشنهادی در فضای رنگی RGB
۵۸٪	۸۱٪	۷۶٪	۸۷٪	روش پیشنهادی در فضای رنگی RGB با استفاده از ویژگی پیشنهادی ۱
۴۵٪	۶۸,۹۳٪	۸۱٪	۶۰٪	روش پیشنهادی در فضای رنگی RGB با استفاده از ویژگی پیشنهادی ۲
۵۹٪	۸۰,۹٪	۸۰٪	۸۲٪	روش پیشنهادی در فضای رنگی HSV
۶۴٪	۸۴,۷٪	۸۰٪	۹۰٪	روش پیشنهادی در فضای رنگی HSV با استفاده از ویژگی پیشنهادی ۱
۶۵٪	۹۱٪	۸۰٪	۹۷٪	روش پیشنهادی در فضای رنگی YCbCr
۵۸٪	۸۳٪	۷۷٪	۹۲٪	روش پیشنهادی در فضای رنگی YCbCr با استفاده از ویژگی پیشنهادی ۱

مقاله [۱۹] بدون CBTF در فضای رنگی RGB	-	-	-	۳۲,۰۴٪
مقاله [۱۹] بدون CBTF در فضای رنگی HSV	-	-	-	۳۴,۴۴٪
مقاله [۱۹] بعد از اعمال CBTF در فضای رنگی RGB	-	-	-	۵۴,۱۶٪
مقاله [۱۹] بعد از اعمال CBTF در فضای رنگی HSV	-	-	-	۲۳,۴۳٪

در این آزمایشات برای بدست آوردن میزان شباهت افراد نسبت به فرد جدید و افراد موجود در پایگاه- داده ها در وضعیت‌های ۳،۱، ۴ و ۶ که از هیستوگرام قسمت‌های مختلف بدن استفاده شده از رابطه (۴-۲۵) استفاده کرده‌ایم که در این رابطه حد‌آستانه را برابر ۰,۷ در نظر گرفته‌ایم و تعداد binها برای آزمایشات را در این وضعیت‌ها ۲۵۶ بوده است در این قسمت هیستوگرام رنگ قسمت‌های میان‌تنه و پایین تنه به صورت جدا در نظر گرفته شده است. برای استفاده از میزان شباهت و شناسایی مجدد با استفاده از تجزیه و تحلیل مولفه‌های اصلی قوی فازی غیر خطی که در وضعیت‌های ۲ و ۵ بوده از رابطه (۴-۳۶) استفاده شده که حد‌آستانه برای این رابطه ۵۵۰ در نظر گرفته شده است و برای بالا بردن سرعت محاسبات تعداد binها با استفاد از این روش ۳۲ بوده که از قسمت میان تنه بدن به صورت جدا استفاده شده است.

۵-۱۰- جمع‌بندی

در این فصل به پیاده‌سازی و توضیح روش‌های مطرح شده در این پایان نامه پرداخته‌ایم در ابتدا با استفاده از مرحله آموزش که می‌تواند به صورت جدا در هنگام راه‌اندازی انجام گیرد پارامترهای لازم برای سیستم‌فازی به منظور تشخیص موقعیت افراد نسبت به دوربین‌ها محاسبه شده و هم زمان تابع شدت روشنایی در هنگام عبور افراد از یک دوربین به دوربین دیگر با استفاده از روش معرفی شده بدست آمده است. پارامترها از فردی که در مرحله آزمون وجود ندارد، بدست آمده که این نکته می- تواند نشان دهنده وابسته نبودن روش پیشنهادی به افراد و محیط مورد آزمایش باشد و در گام بعدی در مرحله آزمون در ابتدا با توجه به اینکه فرد در کدام دوربین قرار دارد از تابع انتقال روشنایی برای برطرف کردن تغییرات شدت روشنایی افراد در هنگام عبور افراد از بین دوربین‌ها استفاده نموده‌ایم و سپس موقعیت فرد نسبت به دوربین بدست آورده شده و با توجه به موقعیت این افراد از ویژگی هر

قسمت به صورت جدا استفاده نموده‌ایم و عمل ردیابی و شناسایی مجدد را انجام داده‌ایم. آزمایشات بر روی مجموعه داده معرفی شده در فضاهاى رنگی متفاوت مورد ارزیابی قرار گرفت که فضای رنگی YCbCr نسبت به فضاهاى رنگی RGB و HSV نتایج بهتری از خود نشان داده است.

فصل ششم

نتیجه گیری و کارهای آتی

۶-۱- نتیجه‌گیری

در این پایان‌نامه، ردیابی افراد در شبکه‌ای از دوربین‌ها با دید مجزا (هیچ نقطه‌ی مشترکی بین زاویه دو دوربین وجود ندارد)، در محیط‌های بسته مانند ادارات و مراکز خرید مورد بررسی قرار گرفته است. ردیابی افراد در این نوع از سیستم‌ها عموماً به دو بخش شناسایی اهداف متحرک و انطباق افراد در دوربین و شناسایی مجدد افراد در بین دوربین‌ها تقسیم می‌شود. در ردیابی در یک دوربین ابتدا باید اهداف متحرک از تصاویر ویدیویی استخراج شوند که کیفیت تصویر استخراج شده از افراد در ردیابی افراد بسیار حائز اهمیت می‌باشد. ردیابی به دلیل وجود سایه، تغییر شرایط نوری و تغییر اندازه و ظاهر با مشکلات بسیاری روبرو می‌شود به همین منظور در این پایان‌نامه به دنبال روشی قدرتمند برای بالابردن سیستم ردیاب انسان در شبکه‌ای از دوربین‌ها می‌باشیم که در برابر این مشکلات مقاوم باشد. انتخاب ویژگی مناسب برای ردیابی که هم از لحاظ سرعت بالا (بار محاسبات پایین) و هم دقت خوبی داشته باشد بسیار با ارزش می‌باشد. به همین منظور ابتدا به بررسی ویژگی‌های که در این نوع از سیستم‌ها برای ردیابی و شناسایی مجدد افراد استفاده شده مورد بررسی قرار داده‌ایم. سیستم‌های که در سال‌های اخیر معرفی شده عموماً از سه نوع ویژگی حرکت جسم، مدل‌های هندسی و ظاهر اشیاء استفاده کرده‌اند. در ردیابی بر پایه حرکت، اگر پارامترهای حرکت شی فریم فعلی مانند مکان، سرعت و شتاب شبیه پارامترهای حرکت شی فریم قبلی باشد هر دو شی به عنوان یک شی شناسایی می‌شود. روش‌های فیلتر کالمن، فیلتر ذرات و جریان‌های نوری از جمله معروف‌ترین روش‌ها در این دسته هستند. ردیابی اشیاء بر پایه حرکت مرکز جرم هندسی اشیاء را مورد توجه قرار می‌دهد. حرکت شی با استفاده از فیلترهای خطی و گاهی اوقات غیرخطی مدل می‌شود. هر شی با استفاده از مدل حرکت ردیابی می‌شود که برچسب‌ها به اشیاء به طور مداوم بر اساس موقعیت واقعی و تخمین زده شده و سرعت اختصاص داده می‌شوند. اگر اشیاء زیادی در نزدیکی یکدیگر جابجا شوند، شی با ویژگی‌های حرکت تخمینی و واقعی نزدیک به عنوان شی یک شی در نظر گرفته می‌شوند. به طور کلی، ویژگی-

های ساده حرکت برای ردیابی جسم به خصوص اگر اشیاء در حال حرکت انسان باشند کافی نیست. حرکت انسان ممکن است که به صورت تصادفی در فریم‌ها متفاوت باشد و فیلترهای حرکت قادر نباشند اشیاء را زمانی که باهم هم‌پوشانی دارند ردیابی نمایند. به خصوص اگر به طور طبیعی یک اشیاء از زاویه دید یک دوربین خارج و دوباره وارد شود.

مدل‌های هندسی از ویژگی‌های شکل برای شناسایی و ردیابی اشیاء استفاده می‌کنند. تشخیص لبه، تطبیق مرز شی، گشتاور، مساحت، اندازه، از جمله ویژگی‌های شی محسوب می‌شوند که برای ردیابی مورد استفاده قرار می‌گیرند. ویژگی‌های اشیاء مانند گشتاورها، لبه، اسکلت، مساحت، محیط و مدل شکل به صورت دوبعدی و سه‌بعدی بیشتر استفاده شده است. الگوریتم‌های که نسبت به اندازه، انتقال و چرخش متغیر نیست‌اند برای ردیابی و انطباق اشیاء می‌توانند استفاده شوند. این ویژگی زمانی که تغییرات اندازه متفاوت وجود داشته‌باشد با مشکلات زیادی روبرو می‌شود بخصوص برای زمانی که از چند دوربین استفاده شده و زاویه دید افراد در ورود و خروج آنها متفاوت باشد.

دسته سوم از ویژگی‌ها که در این پایان‌نامه مورد معرفی قرار گرفته ویژگی‌های مبتنی بر ظاهر افراد می‌باشد. تعداد زیادی از ویژگی‌ها و مدل‌ها برای شناسایی اشیاء و ردیابی وجود دارند اما مدل‌های برپایه ظاهر یک‌بعدی و دوبعدی برای ردیابی رایج هستند. مدل مبتنی بر ظاهر به علت استحکام (قدرت بالا) تشخیص جسم صلب و غیر صلب محبوب‌تر هستند. در این روش مدل اشیاء فریم‌های قبلی با اشیاء فریم فعلی با استفاده از مدل ظاهر رنگ تطبیق داده می‌شود.

اطلاعات ظاهر اشیاء بسیار زیاد در شناسایی و ردیابی استفاده شده است. بسیاری از مقالات که در سال‌های اخیر در زمینه ردیابی اجسام غیرصلب منتشر شده از ویژگی‌های ظاهر استفاده کرده‌اند. رنگ اشیاء به عنوان مهمترین ویژگی برای شناسایی معرفی شده‌است که به همین منظور در این پایان‌نامه از این ویژگی با استفاده از هیستوگرام مورد استفاده قرار گرفته است.

یکی از مهمترین مراحل در سیستم‌های ردیاب در شبکه‌ای از دوربین‌ها، شناسایی مجدد افراد می‌باشد که به مشاهده یک شی در زاویه دید یک دوربین و شناسایی دوباره همان شی در زاویه دید همان دوربین یا دوربین دیگر گفته می‌شود. یک شی ممکن است در زاویه دید دوربین، یک یا چندبار وارد شود. اگر الگوریتم‌ها بتوانند کار شناسایی مجدد یک شی و اطلاعات مربوط به زمان و تعداد مرتبه‌های ورود یا خروج این شی از زاویه دید یک دوربین را بدست بیاورد، این کار می‌توان کمک بسیار مهمی برای شناسای و آنالیز رفتار اشیاء انجام دهد.

برای شناسایی مجدد در شبکه‌ای از دوربین‌ها معمولاً ویژگی حرکت نمی‌تواند در این نوع از محیط‌ها مناسب باشد زیرا نقاط ورود و خروج زیادی وجود دارند این در حالی است که ویژگی حرکت به مکان شی بستگی دارد. این امکان وجود دارد که اشیاء از یک مکان خارج شوند و دوباره در همان صحنه ولی این بار از مکان دیگر وارد شوند. کارهای که براساس حرکت معرفی شده عموماً برای مواقعی مناسب می‌باشند که فاصله بین دوربین‌ها کم بوده و افراد ملزم به حفظ وضعیت حرکتی خود از جمله سرعت و جهت حرکت باشند. در صورتی که افراد بین دو دوربین مکث نموده و یا حرکت آزادانه‌ی دیگر انجام دهند سیستم دچار اختلال در تشخیص می‌گردد. به طور مشابه خواص هندسی نمی‌توانند نتایج خوبی برای شناسایی دوباره اشیاء برای اشیاء غیرصلب به دلیل تغییر اندازه بسیار زیاد (دور و نزدیک شدن شی نسبت به دوربین‌های متفاوت)، از خود نشان دهند.

بعضی از محققان از ویژگی بافت استفاده کرده‌اند که به دلیل کیفیت پایین دوربین‌های نظارتی فقط استفاده از اطلاعات استخراج شده از بافت برای ردیابی صحیح مناسب نیست.

بعضی از کارها، از انطباق نقاط کلیدی برای برقراری ارتباط بین اشیاء استفاده کرده‌اند به مانند SIFT، PCA-SIFT و SURF که این ویژگی‌ها به دلیل بار محاسبات بالا کمتر برای ردیابی در سیستم‌های بلادرنگ استفاده شده است.

از روش‌های دیگر که برای شناسایی مجدد استفاده شده ویژگی‌های بیومتریک استخراج شده از افراد می‌باشد. یکی از این ویژگی، طریقه راه رفتن افراد است. برای استخراج ویژگی‌های راه رفتن، از مدل‌های حرکت برپایه اطلاعات قسمت‌های بیومتریک اشیاء که توصیف کننده زاویه حرکت از زانو و مفصل ران در حالت‌های متفاوت از دوره‌ای از راه رفتن می‌باشد استفاده می‌شود. استفاده از راه رفتن به دلیل نیاز به پردازش نسبتاً سنگینی نیاز و همچنین نمونه‌های ویدئویی در شرایط کاملاً کنترل شده در یک محیط و توسط یک دوربین کمتر برای این نوع از سیستم‌ها استفاده شده است. یکی دیگر از خواص بیومتریک می‌توان به چهره انسان اشاره کرد استفاده از چهره، افراد را ملزم به حرکت در فاصله‌ی کمی از دوربین‌ها می‌نماید تصاویری برای این کار استفاده می‌شوند باید با کیفیت بسیار بالای باشند که مشکلات زیادی را ایجاد می‌نماید.

ویژگی‌های مبتنی بر رنگ به خصوص به دلیل سادگی و سرعت بالا معمولاً در سیستم‌های بلادرنگ بسیار مورد توجه و استفاده قرار گرفته‌اند. ولی یکی از مشکلات اساسی در این نوع از سیستم‌ها تغییر شدت روشنایی در زمان عبور افراد در بین دوربین‌ها می‌باشد که به همین منظور در این پایان‌نامه برای برطرف کردن این تغییر روشنایی از یک تابع انتقال روشنایی با عنوان تابع انتقال روشنایی تجمعی استفاده کرده‌ایم که در بهبود عملکرد و بالا بردن دقت شناسایی تاثیر مناسبی داشته است. و از آنجای که استفاد از هیستوگرام رنگ کل بدن انسان این امکان را ایجاد می‌کنند که تفاوت‌های بین قسمت‌های مختلف بدن (لباس) را به خوبی نشان ندهد به همین منظور از مطالعات کلینیکی استفاده شده، که با استفاده از نسبت‌های موجود در بدن انسان نسبت به کل قد، بدن به سه قسمت سر، میان‌تنه و پایین تنه تقسیم شده است. یکی دیگر از مشکلات در این نوع از سیستم‌ها متفاوت بودن مکان‌های ورود و خروج افراد نسبت به دید دوربین‌ها می‌باشد که این امکان وجود دارد بعضی از ویژگی‌ها در مکان‌ها متفاوت نتایج رضایت بخشی از خود نشان ندهد در همین راستا برای استفاده از هر ویژگی در یک قسمت مجزا از یک سیستم فازی که از پنج ویژگی ارتفاع، عرض، اندازه، فاصله پیکسلی و زاویه

افراد استفاده شده است. سیستم‌فازی محیط تحت پوشش هر دوربین به شش قسمت که در آن سه قسمت رو به دوربین و سه قسمت پشت به دوربین می‌باشد تقسیم شده است. ویژگی دیگری که در این پایان‌نامه معرفی شده تجزیه و تحلیل مولفه‌های اصلی قوی فازی بوده که نتایج خوبی در ردیابی و شناسایی مجدد را نشان داده است.

نتایج بدست آمده با استفاده از ویژگی‌های مطرح شده نتایج خوبی نسبت به کارهای گذشته که بر روی مجموعه داده معرفی شده در این پایان‌نامه انجام گرفته، نشان داده است.

۶-۲- پیشنهادها

کارهای موجود در زمینه‌ی ردیابی در چند دوربین با دید مجزا، از پایگاه داده مخصوص به خود استفاده می‌نمایند. دلیل این امر عدم وجود پایگاه داده استاندارد در این زمینه می‌باشد. به دلیل استفاده از داده‌های مختلف (که از نظر شرایط و میزان پیچیدگی متفاوت می‌باشند) مقایسه کارهای انجام شده را با مشکل مواجه می‌نماید. بنابراین نیاز به پایگاه داده‌ی استاندارد به شدت احساس می‌گردد.

استفاده از دوربین‌های سه‌بعدی می‌تواند کمک بسیاری برای بالا بردن عملکرد این نوع از سیستم‌ها از خود نشان دهد. توانایی سیستم‌ها را برای بدست آوردن ساختار بدن و اسکلت بدن با توجه به اینکه علاوه بر تصویر هر شی، عمق آن را می‌توان استخراج کرد. البته با توجه به اینکه این نوع از دوربین‌ها مشکلاتی در عمق تحت پوشش دارند (تاکنون حداکثر ۳ متر) برای ردیابی محیط‌های پیچیده نمی‌توانند فعلاً مناسب باشند.

استفاده از مفاهیم فازی برای استخراج ویژگی و ردیابی انسان در محیط‌های با چندین دوربین با دید مجزا.

تحلیل و آنالیز رفتار افراد و نوشتن اسکریپت: در بسیاری از مواقع در سیستم ردیابی نیاز نمی‌باشد که تمام مکان‌های که توسط یک فرد مشاهده شده را به صورت یک فیلم نمایش دهد که یک کاربر باید

تمام این مراحل را به صورت دستی مشاهده نماید. در واقع نیاز به سیستمی می‌باشد که رفتار افراد در کل سیستم تحلیل و آنالیز قرار گرفته و در نهایت کارهای خاصی که یک فرد در کل زمان حضور در زاویه دید دوربین‌ها انجام داده به صورت متن (نوشتن اسکرپت) به کاربر سیستم داده شود.

مراجع

- [1] C. Veenman, M. Reinders, and E. Backer, "Resolving motion correspondence for densely moving points," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 1, pp. 54–72, 2001.
- [2] D. Serby, S. Koller-Meier, and L. V. Gool, "Probabilistic object tracking using multiple features," In *IEEE International Conference of Pattern Recognition (ICPR)*, pp. 184–187, 2004.
- [3] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol 25, pp. 564–575, May 2003.
- [4] A. Yilmaz, X. LI, and M. Sham, "Contour based object tracking with occlusion handling in video acquired using mobile cameras," *IEEE Transactionson pattern analysis and machine in intelligence*, vol. 26, no. 11, pp. 1531–1536, 2004.
- [5] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Computer Survey*, vol. 38, no. 4, pp. 1–45, 2006.
- [6] A. Ali, and J. Aggarwal, "Segmentation and recognition of continuous human activity," In *IEEE-Workshop on Detection and Recognition of Events in Video*, pp. 28–35, 2001.
- [7] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capure and analysis," *Computer Vision and Image Understanding*, vol. 104, no. 2, pp. 90–126, 2006.
- [8] M. Enzweiler and D.M. Gavrila, "Monocular pedestrian detection: Survey and experiments," *PAMI*, vol. 31, no. 12, pp. 2179–2195, 2009.
- [9] P. Viola, M. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," *International Journal Computer vision*, vol. 33, no. 2, pp. 153–161, 2005.
- [10] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," In *Computer Vision and Pattern Recognition, CVPR05*, pp. 886–893, 2005.
- [11] C. Wöhler and J. Anlauf, "An adaptable time-delay neural-network algorithm for image sequence analysis," *IEEE Trans. Neural Networks*, vol. 10, no. 6, pp. 1531–1536, 1999.
- [12] D. M. Gavrila and S. Munder, "Multi-cue pedestrian detection and tracking from a moving vehicle," *International Journal Computer vision*, vol. 73, no. 1, pp. 41–59, 2007.
- [13] D. Geronimo, A. M. Lopez, A. D. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," *IEEE Transactions ON Pattern Analysis and Machine Intelligence, (PAMI)*, vol. 32, no. 2, pp. 1239–1258, 2010.
- [14] A.E Elgammal and L.S. Davis, "Probabilistic framework for segmenting people under occlusion," In *IEEE International Conference on Computer Vision, ICCV001*, pp. 145–152, 2001.
- [15] Q. Cai, A. Mitiche, and J.K. Aggarwal, "Tracking human motion in an indoor environment," In *IEEE International Conference on Image Processing, ICIP03*, pp. 215–218, 1995.
- [16] S. Yasutomi and H. Mori, "A method for discriminating of pedestrian based on rhythm," In *International Conference on Intelligent Robots and Systems*, pp. 988-995, 1994.
- [17] L. Q. Xu and D. C. Hogg, "Neural networks in human motion tracking - an experimental study," *Image and Vision Computing*, vol. 15, pp. 607-615, 1997.
- [18] H. Elzein, S. Lakshmanan, and P.Watta, "A motion and shape-based pedestrian detection algorithm," In *IEEE Intelligent Vehicles Symposium*, pp. 500–504, 2003.

- [19] R. E. Kalman, "A new approach to linear filtering and prediction problems," T-ASME, pp. 35–45, 1960.
- [20] R. G. Brown and P. Y. C. Hwang. Introduction to Random Signal and Applied Kalman Filtering. John Wiley & Sons, Inc , USA, 2nd edition, 1992.
- [21] H. Medeiros, J. Park, and A. Kak, "Distributed object tracking using a cluster-based kalman filter in wireless camera networks," IEEE Journal of Selected Topics in Signal Processing, vol. 2, no. 4, pp. 448–463, 2008.
- [22] H. Lee and H. Ko, "Detection of occluded multiple objects using occlusion activity detection and object association," ISPACS, pp. 100–105, 2004.
- [23] A. Doucet, S. Godsill, and C. Andrieu, "On sequential monte carlo sampling methods for bayesian filtering," Statistics and Computing, vol. 10, no. 3, pp. 197–208, 2000.
- [24] J. Czyz, B. Ristic, and B. Macq, "A particle filter for joint detection and tracking of color objects," IVC, vol. 25, no. 8, pp. 1271–1281, 2007.
- [25] J. Martinez-Del-Rincon, C. Orrite-Urunuela, and JE. Herrero-Jaraba, "An efficient particle filter for color-based tracking in complex scenes," In Advanced Video and Signal Based Surveillance AVSS07, pp. 176–181, 2007.
- [26] I. Mikic, M. Trivedi, E. Hunter, and P. Cosman, "Human body model acquisition and tracking using voxel data," International Journal of Computer Vision, vol. 53, no. 3, pp. 199–223, 2003.
- [27] X. Luo and S. Bhandarkar, "Multiple object tracking using elastic matchings," In Advanced Video and Signal based Surveillance, AVSS05, pp. 123–128, 2005.
- [28] M. Hu, "Visual pattern recognition by moment invariants," IRE Transactions on Information Theory, vol. 8, pp. 179–187, 1962.
- [29] A. Dailianas, R.B. Allen, and P. England, "Comparison of automatic video segmentation algorithms," In SPIE Conference on Integration Issues in Large Commercial Media Delivery systems, pp. 2–16, 1995.
- [30] A Kadyrov and M Petrou, "Object descriptors invariant to affine distortions," In British Machine Vision Conference, BMVC01, pp. 391–400, 2001.
- [31] D. Xu and H. Li, "Geometric moment invariants," Pattern Recognition Journal, vol. 41, no. 1, pp. 240–249, 2008.
- [32] M. Holm, "Towards automatic rectification of satellite images using feature based matching," In International Geoscience and Remote Sensing Symposium, pp. 2439–2442, 1991.
- [33] A. Martelli, "Edge detection using heuristic search methods," Computer Graphics and Image Processing, vol. 1, no. 2, pp. 169–182, 1972.
- [34] A. Rosenfeld and A. C. Kak, Digital Picture Processing. Academic Press, USA, 1976.
- [35] R. L. T. Cederberg, "Chain-link coding and segmentation for raster scan devices," Computer Graphics and Image Processing, vol. 10, no. 3, pp. 224–234, 1979.
- [36] G. P. Ashkar and J. W. Modestino, "The contour extraction problem with biomedical applications," Computer Graphics and Image Processing, vol. 7, no. 3, pp. 331–355, 1978.

- [37] J. M. Lester, H. A. Williams, B. A. Weintraub, and J. F. Brenner, "Two graph searching techniques for boundary finding in white blood cell images," *Computers in Biology and Medicine*, vol. 8, no. 4, pp. 293–308, 1978.
- [38] T. Cootes, G. EDWARDS, and C. TAYLOR, "Robust real-time periodic motion detection, analysis and applications," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, PAMI, vol. 23, no. 6, pp. 681–685, 2001.
- [39] S. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, PAMI, vol. 18, no. 9, pp. 884–900, 1996
- [40] N. Parragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, vol. 46, no. 3, pp. 223–247, 2002.
- [41] A.M. Elgammal, R. Duraiswami, and L.S. Davis, "Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking," *PAMI*, vol. 25, no. 11, pp. 1499–1504, 2003.
- [42] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non rigid objects using mean shift," In *CVPR*, pp. 142–149, 2000.
- [43] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer, "Multi-camera multiperson tracking for easy living," In *Proceedings of the Third IEEE International Workshop on Visual Surveillance (VS'2000)*, pp. 3–10, 2000.
- [44] Y. Wei, J. Sun, X. Tang, and H. Yeung, "Interactive offline tracking for color objects," In *ICCV07*, pp. 1–8, 2007.
- [45] C. J. Walder and B. C. Lovell, "Face and object recognition and detection using color vector quantization," In *Fourth Australasian Workshop on Signal Processing and Applications*, pp. 27–30, 2002.
- [46] S. T. Birchfield and S. Rangarajan, "Spatiograms versus histograms for region-based tracking," In *CVPR05*, pp. 1158–1163, 2005.
- [47] A. Mittal and L.S. Davis, "M2tracker: A multi-view approach to segmenting and tracking people in a cluttered scene," *IJCV*, vol. 51, no. 3, pp. 189–203, 2003.
- [48] N. Thome and S. Miguet, "A robust appearance model for tracking human motion," In *Advanced Video and Signal Based Surveillance AVSS*, pp. 528–533, 2005.
- [49] N. Thome, D. Merad, and S. Miguet, "Human body part labeling and tracking using graph matching theory," In *AVSS06*, pp. 38–43, 2006.
- [50] K. Sato and J. K. Aggarwal, "Temporal spatio-velocity transform and its application to tracking and interaction," *CVIU*, vol. 96, no. 2, pp. 100–128, 2004.
- [51] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision (IJCV)*, vol. 60, no. 2, pp. 91–110, 2004.
- [52] H. Bay, A. Ess, T. Tuytelaars, and L.V. Gool, "Surf: Speeded up robust features," *Computer Vision and Image Understanding (CVIU)*, vol. 110, no. 3, pp. 346–359, 2008.
- [53] L. Juan and O. Gwun, "A comparison of sift, pcasift and surf," *International Journal of Image Processing*, IJIP, vol. 3, pp. 143–152, 2009.

- [54] P. Fieguth and D. Terzopoulos, "Color-based tracking of heads and other mobile objects at video frame rates," In CVPR97, pp. 21–27, 1997.
- [55] D. M. Gavrila and V. Philomin, "Real-time object detection for smart vehicles," In ICCV99, pp. 87–93, 1999.
- [56] F. D. Torre, C. Vallespi, P. E. Rybski, M. Veloso, and T. Kanade, "Learning to track multiple people in omni-directional video," In International Conference on Robotics and Automation99, pp. 4150–4155, 2005.
- [57] M. Black and A. Jepson. Eigenttracking, "Robust matching and tracking of articulated objects using a view-based representation," Int. J. Comput. Vision, vol. 26, no. 1, pp. 63–84, 1998.
- [58] B. Mughadam and A. Pentland, "Probabilistic visual learning for object representation," PAMI, vol. 9, no. 7, pp. 696–710, 1997.
- [59] A. Alahi, P. Vandergheynst, M. Bierlaire, and M. Kunt, "Cascade of descriptors to detect and track objects across any network of cameras," CVIU, vol. 114, no. 6, pp. 624–640, 2010.
- [60] N. Noceti, A. Destrero, A. Lovato, and F. Odone, "Combined motion and appearance models for robust object tracking in real-time," AVSS09, pp. 412–417, 2009.
- [61] S. M. Khan and M. Shah, "Tracking multiple occluding people by localizing on multiple scene planes," PAMI, vol. 31, no. 3, pp. 505–519, 2009.
- [62] K. Kim and L. Davis, "Multi-camera tracking and segmentation of occluded people on ground plane using search-guided particle filtering," In Ninth European Conference Computer Vision, 2006.
- [63] O. Javed, K. Shafique, Z. Rasheed, and M. Shah, "Modeling inter-camera space-time and appearance relationships for tracking across non-overlapping views," Computer Vision and Image Understanding, vol. 109, no. 2, pp. 146–162, 2008.
- [64] A. Ilyas, Object Tracking and Re-Identification In Multi-Camera Environments, phd Thesis, 2011.
- [65] D. Makris, T. J. Ellis, and J. K. Black, "Bridging the gaps between cameras," In IEEE Conference on Computer Vision and Pattern Recognition, 2004.
- [66] A. Rahimi and T. Darrell, "Simultaneous calibration and tracking with a network of nonoverlapping sensors," In IEEE Conference on Computer Vision and Pattern Recognition, 2004.
- [67] L. J. Zhu, J. N. Hwang, and H. Y. Cheng, "Tracking of multiple objects across multiple cameras with overlapping and non-overlapping vie," In IEEE International Symposium on Circuits and Systems, (IS-CAS), pp. 1056–1060, 2009.
- [68] O. Javed, K. Shafique, and M. Shah, "Appearance modeling for tracking in multiple nonoverlapping cameras," In IEEE Conference on Computer Vision and Pattern Recognition, pp. 26–33, 2005.
- [69] T. Pham, M. Worring, and A. Smeulders, "A multi-camera visual surveillance system for tracking re-occurrences of people," In International Conference on on Smart Distributed Cameras, 2007.
- [70] B. Prosser, S.G. Gong, and T. Xiang, "Multi-camera matching using bi-directional cumulative brightness transfer functions," In BMVC08, 2008.
- [71] T. D. Orazio, P. L. Mazzeo, and P. Spagnolo, "Color brightness trasfer function evaluation for non overlapping multi camera tracking," In International conference on distributed cameras (ICDS09), pp. 1–6, 2009.

- [72] F. Porikli and A. Divakaran, "Multi-camera calibration, object tracking and query generation," In ICME 03, pp. 653–656, 2003.
- [73] V. Kettner and R. Zabih, "Bayesian multi-camera surveillance," In Proceedings of the Computer Vision and Pattern Recognition (CVPR), pp. 253–259, 1999.
- [74] M. Lantagne, M. Parizeau, and R. Bergevin, "Vip: Vision tool for comparing images of people," In IEEE Conference on Vision Interface, pp. 35–42, 2003.
- [75] L. Vacchetti, V. Lepetit, and P. Fua, "Combining edge and texture information for realtime accurate 3d camera tracking," In Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality, pp. 48–57, 2004.
- [76] C. Arth, C. Leistner, and H. Bischof, "Object reacquisition and tracking in large-scale smart camera networks," In IEEE International Conference on Smart Distributed Cameras, 2007.
- [77] N. Gheissari, T. Sebastian, and R. Hartley, "Person reidentification using spatiotemporal appearance," In International Conference on Computer Vision and Pattern Recognition, (CVPR), pp. 1528–1535, 2006.
- [78] Y. Ke and R. Sukthankar, "Pca-sift: A more distinctive representation for local image descriptors," In Proceedings of the Computer Vision and Pattern Recognition (CVPR), pp. 506–513, 2004.
- [79] O. Hamdoun, F. Moutarde, B. Stanculescu, and B. Steux, "Person re-identification in multi-camera system by signature based on interest point descriptors collected on short video sequences," In ICDSC08, pp. 1–6, 2008.
- [80] G. Johansson, "Visual perception of biological motion and a model for its analysis," Perception and Psychophysics, vol. 14, pp. 201–211, 1973.
- [81] M. P. Murray, "Gait as a total pattern of movement," American Journal of Physical Medicine, vol. 46, no. 1, pp. 290–333, 1967.
- [82] P. S. Huang, C. J. Harris, and M. S. Nixon, "Human gait recognition in canonical space using temporal templates," IEEE Proceedings of Vision, Image and Signal Processing, vol. 146, no. 2, pp. 93–100, 1999.
- [83] P. M. Roth, H. Grabner, D. Skocaj, H. Bischof, and A. Leonardis, "On-line conservative learning for person detection," In IEEE Workshop on VS-PETS, pages 223–230, 2005.
- [84] I. Bouchrika, J. N. Carter, and M. S. Nixon, "Recognizing people in non-intersecting camera views," In 3rd International Conference on Crime Detection and Prevention (ICDP), pp. 1–6, 2009.
- [85] C. BenAbdelkader, R. Cutler, and L. Davis, "Person identification using automatic height and stride estimation," Pattern Recognition, 2002, Proceedings. 16th International Conference on, IEEE, p. 377–380, 2005.
- [86] T. Darrel, G. Gordon, M. Harvill, and J. Woodfill, "Integrated person tracking using stereo, color, and pattern detection," International Journal of Computer Vision, vol. 37, p. 175–185, 2000.
- [87] M. H Tan and S. Ranganath, "Multi-camera people tracking using Bayesian networks," Proceedings of the 2003 joint Conference of the Fourth International Conference on Information Communications and signal Processing and the Fourth Pacific Tim Conference on Multimedia, IEEE Press, pp. 1335–1340, 2003.

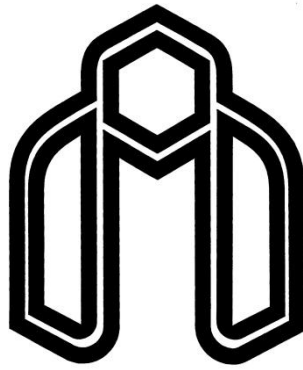
- [88] S. Bak, E. Corvee, F. Brémond, and M. Thonnat, "Person re-identification using spatial covariance regions of human body parts," In 7th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS), 2010.
- [89] C. Madden, Tracking People Tracking Across Disjoint Camera Views, phd Thesis, 2009.
- [90] D. Rowe, Reviewing Detections and tracking Approaches, February 2008
- [91] D. R. Magee, "Tracking multiple vehicles using foreground, background and motion models," Image and Vision Computing, vol. 22, pp. 143–155, 2004.
- [92] S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler, "Tracking Groups of People," Computer Vision and Image Understanding, vol. 80, no. 1, pp. 42-56, 2000.
- [93] S. M. Yoon, and H. Kim, "Real-time multiple people detection using skin color, motion and appearance information," Proc. 2004 IEEE Int. Workshop on Robot and Human Interactive Communication, Kurashiki, Okayama, Japan, pp. 331-334, September 20-22, (2004).
- [94] S. Haritaoglu, D. Harwood, and L.S. Davis, "W4: Real-Time Surveillance of People and Their Activities," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 8, Aug. (2000).
- [95] C. Stauffer and W. Grimson, "Adaptive background mixture models for realtime tracking," Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 2, pp. 246-252, (1999).
- [96] T. Ellis, "multi camera video surveillance," Proc. 36th Annual 2002 Int. Carnahan Conf. on Security Technology, Atlantic City, New Jersey, pp.228-233, (2002).
- [97] L. Gui, C. Garboni, P. Horvath and P. Covacs, Motion Tracking, SSIP 2003.
- [98] F. Remondino, A. Roditakis, "Human Figure Reconstruction and Modeling from Single Image or Monocular Video Sequence," 4th International Conference on 3D Digital Imaging and Modeling (3DIM), October 2003.
- [99] F. Remondino, A. Roditakis, "3D Reconstruction of Human Skeleton from Single Images or Monocular Video Sequences," 25th Pattern Recognition Symposium, Lecture Notes in Computer Science, Springer, DAGM 03, pp. 100-107, 2003.
- [100] k. Jeon, Object Matching in Disjoint Cameras Using a Color Transfer Approach, 2006.
- [101] C. Madden, E.D. Cheng, and M. Piccardi, "Tracking people across disjoint camera views by an illumination-tolerant appearance representation," Machine Vision and Application, vol. 18, p.233-247, Mar, 2007.
- [102] P. L. Mazzeo, L. Giove, G. M. Moramarco, P. Spagnolo, and M. Leo, "HSV and RGB color histograms comparing for objects tracking among non-overlapping FOVs, using CBTF," Workshop on Activity monitoring by multi-camera surveillance systems (AMMCSS) in conjunction with 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance, 2011.
- [103] C. M Bishop, Pattern Recognition and Machine Learning. Springer, 2006.
- [104] P. Luukka, "A New Nonlinear Fuzzy Robust PCA Algorithm and Similarity Classifier in Classification of Medical Data Sets," International Journal of Fuzzy Systems, vol. 13, no. 3, September 2011.
- [105] T. N. Yang and S. D. Wang, "Robust algorithms for principal component analysis," Pattern Recognition Letters, vol. 20, pp. 927-933, 1999.
- [106] E. E. Oja, "The Nonlinear PCA Learning Rule and Signal Separation-Mathematical Analysis," Technical Report, Helsinki University of Technology, pp. 26, 1995.

Abstract

Analysis, motions and behavior of human have always been interesting subjects in different areas of science. The aim of designing such systems is to get better understanding on the behavior of human using video images sequence. Human behavior analysis can be assessed from various points of views such as initialization, tracking as well as pose recovery and recognition. Tracking is by far one of the most significant steps in behavior analysis conducted using network covering a wide range of areas. To reduce the amount of cost in such cases, cameras with disjoint views are often used. However, features like appearance of person, environmental illumination, location of person respect to planted cameras as well as camera's angle of view involved in such systems make the results unreliable along with considerable errors.

In this thesis, an approach was introduced to be used in a network of cameras with disjoint views. The tracking process was performed in identification and tracking steps in the view of one cameras and matching with other cameras. After doing necessary preprocessing steps and elimination of useless parts, three global features including motion, geometrical model and appearance were evaluated aiming to present their weakness and strength points. Since methods working based on motions and geometrical models were not able to provide reliable results, color histogram was used as the best feature extracted from persons appearances. Color histogram is a suitable method for such systems as it is a very simple but effective approach. However, it cannot be properly used once brightness and size of the people going through the cameras are changed. Thus, cumulative brightness transform function was used to reduce the variation of brightness of those persons going through the camera's view. To do tracking process, according to the relative distance of various parts of human body, the body was partitioned into three portions comprising head, torso and bottom part. Since this method cannot be affected by people size, the body can be very well portioned once people get close to the cameras or vice versa. Another method used in the present study to extract feature was nonlinear fuzzy robust principle component analysis (nfrPCA). Those features extracted by this method cannot be always used in entrance and exit of people as they present distinct results in various environmental conditions. Thus, fuzzy system was used to discriminate the area under consideration to three parts of close, middle and far distance to the cameras. To evaluate the efficiency of proposed system, experiments were conducted on a set of data extracted from Italian experiments. The results indicate a better performance for proposed method in YCbCr color space compared to other color space used in previous studies.

Keywords: Human tracking, Network of cameras, Disjoint views, Appearance model, Histogram, Relative distance of human body, Fuzzy system, Cumulative brightness transform function, nonlinear fuzzy robust principle component analysis.



Shahrood University of Technology
Faculty of Computer Engineering

Feature Extraction for Tracking People Across Multiple Camera with Disjoint views for Indoor Environment

Thesis

Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Science (M.Sc)

Ali GHanbari Sorkhi

Supervisors

Dr. Hamid Hassanpour

Associate supervisor

Dr. Morteza Zahedi

Date: 2012