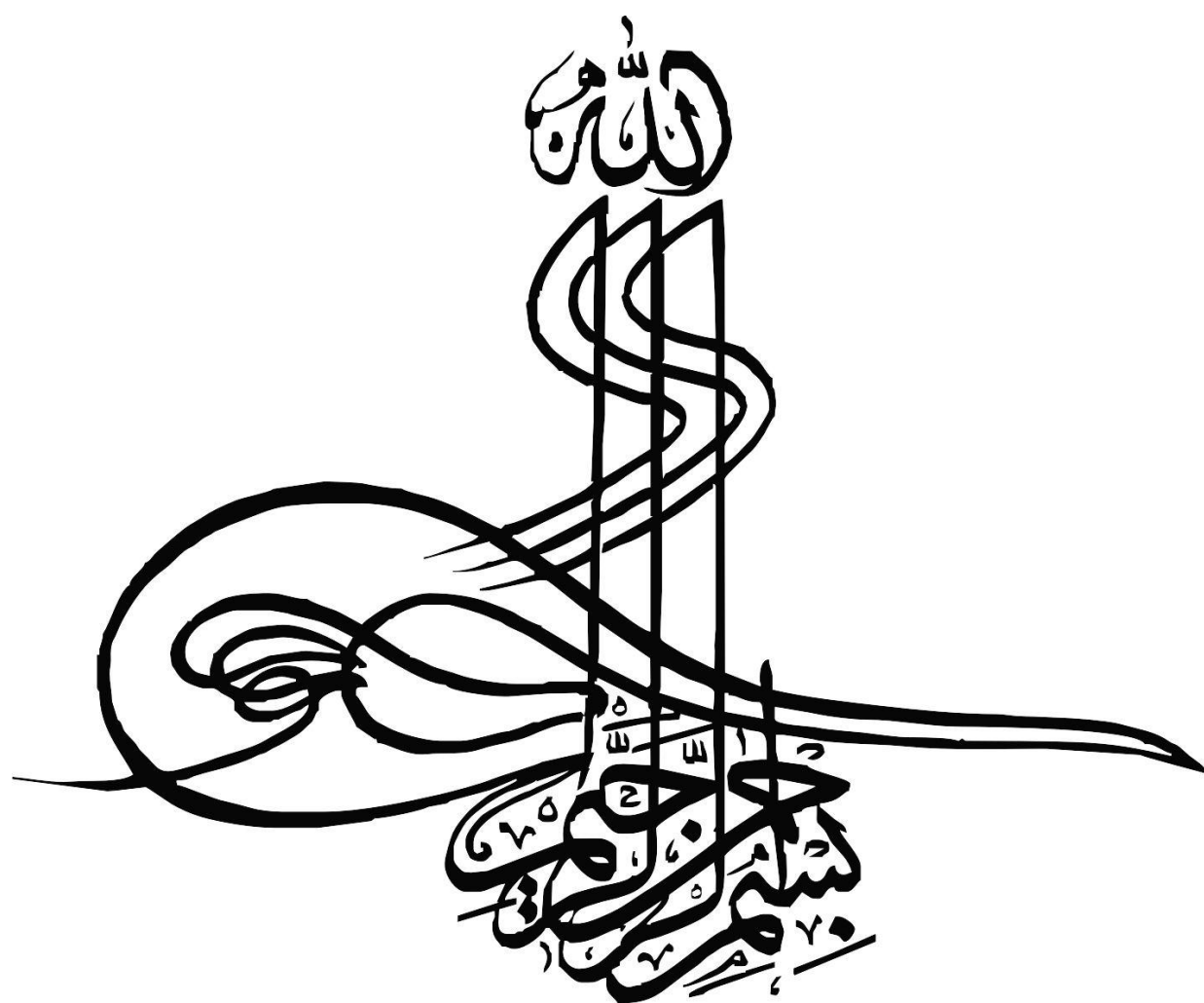






دانشگاه صنعتی شاهرود

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

پایان نامه دوره‌ی کارشناسی ارشد مهندسی کامپیوتر - هوش مصنوعی

ردیابی افراد در شبکه‌ای از دوربین‌ها با دیدهای مجزا در محیط‌های بسته

رحمان یوسف زاده

استاد راهنما:

دکتر حمید حسن پور

استاد مشاور:

دکتر علی اکبر پویان

بهمن ۱۳۸۹

دانشگاه صنعتی شاهرود

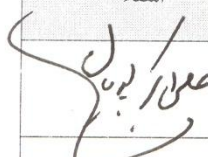
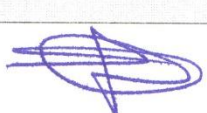
دانشکده: فناوری اطلاعات و کامپیوتر

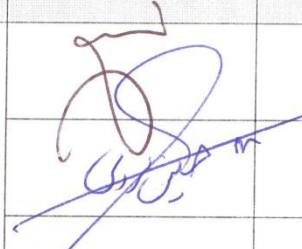
پایان نامه ارشد آقای رحمان یوسف زاده

تحت عنوان:

ردگیری افراد با استفاده از چند دوربین با دیدهای متفاوت در محیط های بسته
در تاریخ ۸۹/۱۱/۱۶ توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد مورد ارزیابی و با

درجه عالی مورد پذیرش قرار گرفت.

امضاء	اساتید مشاور:	امضاء	اساتید راهنما:
	نام و نام خانوادگی: دکتر علی اکبر پویان		نام و نام خانوادگی: دکتر حمید حسن پور
	نام و نام خانوادگی :		نام و نام خانوادگی:

امضاء	نماینده تحصیلات تکمیلی	امضاء	اساتید داور:
	نام و نام خانوادگی : مهندس حامد رحیم اف		نام و نام خانوادگی: دکتر مرتضی زاهدی
			نام و نام خانوادگی: دکتر مروی
			نام و نام خانوادگی:
			نام و نام خانوادگی:

تقدیم بہ

پدر و مادر کو تقدیرم،

بہ پاس تمام فداکاری ما و قلب سرشار از عشق و مہربانیشان کہ ہموارہ آرام بخش بخصات
سخت زندگی و پشتوانہ، حامی و مشوق من بودہ و ہستند،

و برادر و خواہر عزیزم

کہ حس حضورشان، شادی بخش زندگی ام بودہ است.

تشکر و قدردانی

نخست، خداوند بزرگ را شاکرم که یادش آرام بخش جهانیان است. . این تحقیق حاصل تلاشی است که توسط افراد بسیاری حمایت شده است؛ بر خود لازم می‌دانم تشکرات عمیق قلبی‌ام را به آنها تقدیم نمایم. در ابتدا از استاد راهنمای خود، جناب آقای دکتر حمید حسن‌پور کمال تشکر و قدردانی را دارم. بدون راهنمایی و حمایت‌های همه جانبه‌ی ایشان این تحقیق هرگز انجام نمی‌گرفت. همچنین از استاد مشاور خود جناب آقای دکتر پویان تشکر می‌نمایم. نصایح و اندرزهای ایشان همواره باعث دلگرمی و امیدواری بنده بوده است. از جناب آقای دکتر زاهدی که افتخار شاگردی کلاس درس آنها نصیب اینجانب گردید، تشکر می‌نمایم.

حضور بنده در این مقطع، حاصل تلاش معلمان و اساتید بسیاری بوده است که راهنمایی‌های ایشان روشنگر راه بنده بوده است. از تمام اساتید گرامی خود در سالهای دبستان، راهنمایی و دبیرستان از جمله جناب آقای تنها، جناب آقای نجیبی، جناب آقای شیخ بابایی و جناب آقای رحمانی صمیمانه متشکرم. همچنین از تمام اساتید گرامی‌ام در دوره‌ی کاردانی و کارشناسی به خصوص جناب آقای دکتر خانتیموری، جناب آقای مهندس احمدی، سرکار خانم مهندس صفری و جناب آقای مهندس محمدپور تشکر ویژه ای دارم.

در طول این دوره افرادی بسیاری بوده‌اند که همواره نسبت به بنده لطف داشته و سعی در حل مشکلات بنده داشته‌اند. ابتدا بر خود لازم می‌دانم از دوستان خوبم در مرکز کامپیوتر دانشگاه صنعتی شاهرود تشکر نمایم. همچنین از زحمات مسئولین آموزش دانشکده سرکار خانم سرائی و جناب آقای نصرتی قدردانی می‌نمایم که در این مدت همواره با صبر و مهربانی خود، راهنمای بنده در مسائل آموزشی بوده‌اند. از تمام دوستانی که در این مدت، افتخار آشنایشان نصیب اینجانب گردید، به خصوص دوستان هم‌خانه‌ام که باعث گردیدند تا خاطرات به یادماندنی‌ای از این دوره در ذهن اینجانب نقش گیرد تشکر ویژه ای نمایم.

صمیمانه‌ترین سپاسگزاری خود را به خانواده‌ام تقدیم می‌نمایم. کلمه‌ای را نمی‌توان یافت که بتواند فداکاری و توجه آنها را توصیف نماید. از خانواده‌ام بسیار متشکرم و این مجموعه را به آنها تقدیم می‌کنم.

رحمان یوسف زاده، بهمن ۱۳۸۹

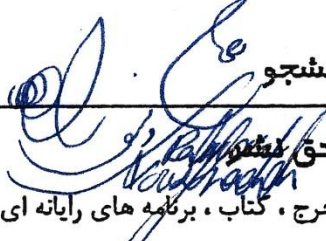
تعهد نامه

اینجانب دکتر محسن باقری دانشجوی دوره کارشناسی ارشد رشته کامپوزیت - هوش مصنوعی
دانشکده کامپوزیت دانشگاه صنعتی شاهرود نویسنده پایان نامه در بابی با فرد در شبدرای از
دو برین ها بابر های خزا در حیطه کمالی تحت راهنمایی دکتر محسن باقری متعهد می شوم .

- 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 می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به سوزنه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ ۱۶/۱۱/۹۸

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 نمی باشد .

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

چکیده

تجزیه و تحلیل رفتار افراد کاربردهای زیادی در حوزه‌های مختلف دارد که از آن جمله می‌توان به بررسی رفتار افراد در مراکز خرید و نیز حفظ امنیت اماکنی مانند بانک‌ها و فرودگاه‌ها اشاره نمود. ردیابی افراد اولین مرحله در این سیستم‌ها می‌باشد. به دلیل محدودیت وسعت دید دوربین، بررسی تمام محیط مورد نظر با استفاده از یک دوربین، در بسیاری از کاربردها امکان پذیر نمی‌باشد. بر این اساس معمولاً شبکه‌ای از دوربین‌ها مورد نیاز است. علاوه بر مسائل مطرح در ردیابی با یک دوربین، ردیابی افراد در شبکه‌ی توزیع شده‌ای از دوربین‌ها، از جهات دیگر نیز چالش بر انگیز می‌باشد. به عنوان مثال ظاهر افراد در دوربین‌های مختلف، به دلیل عواملی از قبیل شرایط نوری محیط، موقعیت افراد نسبت به دوربین، زاویه ی دید و پارامترهای دوربین متفاوت می‌باشد. همچنین حضور افراد در زمان و مکان‌های مختلف در مقابل دید دوربین‌ها تابع قانون خاصی نبوده و افراد می‌توانند به صورت غیر یکنواخت در زمان‌های مختلف، در مکان‌های متفاوتی قرار گیرند. علاوه بر این ممکن است دوربین‌های متوالی فضای دید مشترکی نداشته باشند که این موضوع بر پیچیدگی مسئله می‌افزاید.

در این پایان نامه راهکاری برای ردیابی افراد در شبکه‌ای از دوربین‌ها بدون فضای دید مشترک، ارائه شده است. در این روش، ردیابی افراد طی دو مرحله انجام می‌گیرد. در مرحله اول در هر یک از دوربین‌ها افراد بطور مستقل شناسایی و ردیابی میشوند. در این مرحله اتفاقات رخ داده در هر محیط (دوربین) بر اساس یک ساختمان داده‌ی استاندارد در پایگاه داده‌ی مرکزی ذخیره می‌گردد. در مرحله دوم، تک تک افراد با توجه به اطلاعات ثبت شده در پایگاه داده در کل محیط تحت پوش ردیابی می‌شوند.

به منظور تشخیص افراد در یک دوربین از مدل سازی پس زمینه و محاسبه ی تفاضل فریم فعلی و فریم پس زمینه استفاده شده است. پس از اعمال پیش پردازش‌هایی برای بهبود کیفیت تصویر اشیاء تشخیص داده شده، و حذف قسمت های زائد مانند سایه، خط سیر افراد با اعمال فیلتر کالمن به

دست می‌آید. الگوریتم ارائه شده، با استفاده از اطلاعات حرکتی، قادر به ردیابی افراد در شرایط نوری مختلف، و همچنین هنگام بروز هم پوشانی های معمول بین افراد می‌باشد.

برای برقراری ارتباط بین مشاهدات دریافت شده در مجموعه‌ی دوربین‌ها جهت ردیابی افراد، ویژگی‌های استخراج شده هر شی با ویژگی‌های ذخیره شده در پایگاه داده مقایسه شده و در صورت منطبق بودن برچسب آن بروز رسانی می‌گردد. در این پایان نامه ویژگی جدیدی بر مبنای بافت رنگی تصویر معرفی شده است که نسبت به تغییرات نور دارای ثبات بیشتری بوده و از قابلیت تفکیک بالاتری برخوردار است. این ویژگی از استخراج عناصر واقع بر قطر اصلی ماتریس هم رخداد شی در فضای رنگی YCbCr به دست می‌آید.

به منظور ارزیابی عملکرد الگوریتم ارائه شده، این روش بر روی چند مجموعه‌ی ویدئویی از جمله ویدئوهای ضبط شده از پنج صحنه مختلف از یک محیط کاری که دارای پیچیدگی بالایی می‌باشد، اعمال شده است. نتایج نشان می‌دهد ویژگی ارائه شده دارای دقت بالاتری نسبت به هیستوگرام رنگ و روشهای مشتق شده از آن دارد.

کلمات کلیدی: ردیابی افراد، ماتریس هم رخداد، مدل ظاهر ، فضای رنگ، سیستم های نظارتی

فهرست مطالب

فصل ۱: مقدمه	۱۵
۱-۱- اهداف پایان نامه	۱۹
۲-۱- محدودیت ها و فرضیات مساله	۲۰
۳-۱- کاربردها	۲۱
۴-۱- ساختار پایان نامه	۲۳
فصل ۲: آشکار سازی افراد	۲۴
۱-۲- مقدمه	۲۵
۲-۲- شارنوری	۲۶
۳-۲- تفاضل فریم های متوالی	۲۶
۴-۲- تفاضل زمینه	۲۷
۵-۲- تخمین پس زمینه به کمک مدل ترکیبی گاوسی	۲۸
۱-۵-۲- تئوری مدل گاوسی	۳۰
۲-۵-۲- به روزرسانی مدل ترکیبی گاوسی	۳۱
۳-۵-۲- آشکار سازی پیش زمینه	۳۲
۶-۲- حذف نویز	۳۴
۷-۲- حذف سایه	۳۴
۸-۲- شمای کلی زیر سیستم آشکار سازی افراد	۳۷
۹-۲- نتیجه گیری	۳۸
فصل ۳: ردیابی افراد در یک دوربین	۳۹
۱-۳- مقدمه	۴۰
۲-۳- مروری بر کارهای گذشته	۴۰

۴۱	۱-۲-۳- نحوه‌ی انتخاب معیار ردیابی.....
۴۶	۲-۲-۳- نحوه‌ی تطابق مشاهدات.....
۴۸	۳-۳- فیلترینگ تصادفی.....
۵۰	۱-۳-۳- فیلترینگ بازگشتی بیزین غیر خطی.....
۵۳	۴-۳- فیلتر کالمن.....
۵۹	۵-۳- ردیابی اهداف چندگانه.....
۶۰	۱-۵-۳- فرمول بندی بیزین عمومی برای ردیابی اهداف چندگانه.....
۶۲	۳-۵-۲- ردیابی فرضیات چندگانه (MHT).....
۶۳	۳-۵-۳- الگوریتم JPDAF.....
۶۵	۴-۵-۳- سایر روش های ردیابی.....
۶۷	۶-۳- سیستم پیاده سازی شده.....
۶۸	نتیجه گیری.....
۶۹	فصل ۴: ردیابی اهداف در چند دوربین.....
۷۰	۱-۴- مقدمه.....
۷۳	۲-۴- تعریف مساله.....
۷۵	۳-۴- ویژگی های قابل استفاده برای برقراری ارتباط بین مشاهدات.....
۷۵	۱-۳-۴- رنگ.....
۷۷	۲-۳-۴- ویژگی های مستقل از تغییر مقیاس.....
۸۱	۴-۴- معیار های اندازه گیری شباهت.....
۸۳	۵-۴- ویژگی معرفی شده در پایان نامه.....
۸۴	۱-۵-۴- ماتریس هم‌رخداد.....
۸۴	۶-۴- استخراج ویژگی.....
۸۸	۷-۴- بهبود کارایی با تقسیم ظاهر افراد به چند قسمت.....
۸۹	۸-۴- تطبیق ویژگی ها.....

۹-۴- نتیجه گیری	۸۹
فصل ۵: پیاده سازی سیستم و نتایج حاصله	۹۰
۱-۵- توصیف کلی سیستم پیاده سازی شده	۹۱
۱-۱-۵- زیر سیستم تشخیص و جدا سازی افراد	۹۲
۲-۱-۵- زیر سیستم ردیابی افراد در یک دوربین	۹۲
۳-۱-۵- زیر سیستم برقراری ارتباط بین مشاهدات و ردیابی در چند دوربین	۹۴
۲-۵- مجموعه داده	۹۵
۳-۵- بررسی پارامتر های تاثیرگذار	۱۰۴
۴-۵- نتیجه گیری	۱۰۷
فصل ۶: نتیجه گیری و پیشنهادات برای کارهای آینده	۱۰۸
۱-۶- نتیجه گیری	۱۰۹
۲-۶- پیشنهادات	۱۱۲

فهرست شکل؛

- شکل ۱-۱ ساختار سیستم ردیابی پیشنهاد شده و زیر سیستم های آن ۱۹
- شکل ۱-۲ (الف) چگالی مربوط به یک پیکسل در دویست و پنجاه فریم متوالی (ب) تویع های گاوسی تشخیص داده شده برای مقادیر موجود [۹] ۳۰
- شکل ۲-۲ فریم فعلی ۳۳
- شکل ۳-۲ تصویر زمینه، مدل زمینه با این تصویر برابر نبوده و مجموعه ای از مدل های گاوسی برای هر پیکسل می باشد .. ۳۳
- شکل ۴-۲ نتیجه ی تفاضل فریم فعلی و مدل زمینه ۳۳
- شکل ۵-۲ دو فریم و سایه ی تشخیص داده شده، قسمت های خاکستری نشان دهنده ی سایه می باشند. بخش الف و پ تصاویر اصلی و بخش ب و ت سایه های تشخیص داده شده هستند. ۳۶
- شکل ۶-۲: زیرسیستم آشکار سازی اهداف ۳۷
- شکل ۱-۳ عوامل محدودکننده در ردیابی قطعی افراد (الف) مجاورت (ب) حداکثر سرعت (پ) تغییرات جزئی سرعت (ت) حرکات مشترک (ث) ثبات شکل ۴۷
- شکل ۲-۳ مراحل پیش بینی و بروز رسانی در تئوری بیزین ۵۲
- شکل ۳-۳ مثال ارتباط اطلاعات. (الف) دو هدف (دایره ها) به همراه دو داده ی مشاهده شده (مثلث ها)، جهت ایجاد ارتباط. (ب) حالت ممکن اول در مسئله ی ارتباط اطلاعات. (ج) حالت ممکن دوم برای مسئله ی ارتباط اطلاعات. ۵۹
- شکل ۱-۴ : نحوه ی ایجاد هرم تفاضل گاوسی ها [۴۷] ۷۸
- شکل ۲-۴ نحوه ی تولید توصیف کننده برای نقاط کلیدی در ویژگی مستقل از مقیاس (الف) گرادیان در همسایگی نقطه ی کلیدی (ب) توصیف کننده ۷۹
- شکل ۳-۴ تصاویری با بافت های مختلف و ماتریس هم رخداد متناظر آنها ۸۶
- شکل ۴-۴ تصویر دو فرد و بردار های به دست آمده با استفاده از روش پیشنهادی ۸۷
- شکل ۵-۴ محدوده ی تقسیم ظاهر افراد ۸۸
- شکل ۱-۵ ساختار سیستم ردیابی پیشنهاد شده و زیر سیستم های آن ۹۱
- شکل ۲-۵ زیرسیستم آشکار سازی اهداف ۹۳
- شکل ۴-۵ : نمای دوربین های نصب شده در ساختمان برای جمع آوری پایگاه داده (الف) دوربین ۱، طبقه ی ۴ (ب) دوربین ۲ طبقه ی ۴ (پ) دوربین ۳، طبقه ی ۳ (ت) دوربین ۴، طبقه ی ۳ (ث) دوربین ۵، طبقه ی ۲ ۹۶
- شکل ۵-۵ نمونه ای از ردیابی یک فرد در مجموعه ی دوربین ها از شماره ۱ تا شماره ۴، شروع از راست به چپ ۱۰۳
- شکل ۶-۵ دقت هیستوگرام رنگ و ویژگی پیشنهاد شده در فضاهای رنگی مختلف ۱۰۵
- شکل ۷-۵ : نتایج به دست آمده استفاده از معیار های شباهت بر روی ویژگی تعریف شده ۱۰۵
- شکل ۸-۵ تاثیر ابعاد ماتریس هم جوارى در کارایی سیستم ۱۰۶
- شکل ۹-۵ تاثیر تعداد قطر های اصلی بر کارایی سیستم ۱۰۷

فهرست جداول

- جدول ۱-۵ تصاویر به دست آمده از افراد در هر یک از دوربین ها..... ۹۸
- جدول ۲-۵ مقایسه ی تصاویر مربوط به فرد ۲۱ و با تصاویر موجود در پایگاه داده و نتایج حاصله..... ۱۰۰
- جدول ۳-۵ مقایسه ی تصاویر مربوط به فرد ۳ و ۴ با تصاویر موجود در پایگاه داده و نتایج حاصله..... ۱۰۱
- جدول ۴-۵ مقایسه ی تصاویر مربوط به فرد ۵ با تصاویر موجود در پایگاه داده و نتایج حاصله..... ۱۰۲

فهرست علائم و اصطلاحات اختصاری

ANN	Artificial Neural Network
CAD	Computer Aided Design
DOG	Difference Of Gaussian
EM	Expectation Maximization
ETISEO	Evaluation du Traitement et de l'Interpretation de Sequences Video - Evaluation for video understanding
GMM	Gaussian Mixture Model
HMM	Hidden Markov Model
HSV	Hue, Saturation, Value
JPDAF	Joint Probabilistic Data Association Filter
MHT	Multiple Hypothesis Tracking
MTT	Multi Target Tracking
PDA	Probabilistic Data Association
PDF	Probability Density Function
PETS	Performance Evaluation of Tracking Systems
PTZ	Pan Tilt Zoom
RAG	Region Adjacent Graph
RGB	Red, Green, Blue
ROI	Region Of Interest
SIFT	Scale Invariant Feature Transform
XML	Extensible Markup Language
YCbCr	Luminance, Blue Chrominance, Red Chrominance

فصل ۱: مقدمه

ردیابی افراد از ضروریات سیستم‌های نظارتی و محیط‌های هوشمند می‌باشد که به دلیل کاربرد گسترده در حوزه‌های مختلف، مورد توجه بسیاری از محققان قرار گرفته است. در این فصل مبانی ردیابی اهداف به همراه ضروریات، کاربردها و ملزومات سیستم‌های ردیابی اهداف در چند دوربین با دیدهای مجزا، مورد بررسی قرار خواهد گرفت.

ردیابی افراد از ضروریات سیستم‌های نظارتی و محیط‌های هوشمند می‌باشد که به دلیل کاربرد گسترده در حوزه‌های مختلف، مورد توجه بسیاری از محققان قرار گرفته است. به عنوان مثالی از کاربردهای ردیابی افراد می‌توان به سیستم‌های آنالیز رفتار افراد در مراکز خرید جهت تشخیص رفتار مشکوک و یا ردیابی افراد مجرم در فرودگاه‌ها اشاره نمود. در سالهای اخیر کاربرد ردیابی افراد به طور گسترده‌ای افزایش یافته و علاوه بر مسائل امنیتی، در حوزه‌هایی نظیر اقتصاد و مدیریت نیز مورد استفاده قرار گرفته است. به عنوان مثال مراکز خرید با استفاده از اطلاعات به دست آمده توسط ردیابی مشتریان و تکنیک‌های داده‌کاوی، الگوهای رفتاری مشتریان نسبت به نحوه‌ی چینش محصولات، مسیرهای رفت‌وآمد، ساعات خرید و سایر اطلاعات لازم را استخراج نموده و اقدامات لازم را نسبت به افزایش بهره‌وری اعمال می‌نمایند.

مسئله‌ی ردیابی اهداف در ویدئو را می‌توان به صورت مسئله‌ی تخمین خط سیر^۱ یک شی هنگام عبور شی از محیط تعریف نمود [۱]. به عبارت ساده‌تر هدف از ردیابی افراد مشخص نمودن مکان یک شی در هر لحظه از زمان است. به این منظور، ویژگی‌های سطح پایین تصویر مانند لبه، رنگ و بافت در یک قاب ویدئویی^۲ استخراج شده و در قاب بعدی برای یافتن محل جدید شی مورد استفاده قرار می‌گیرند. در صورتی که افراد به صورت پیوسته قابل مشاهده باشند و شکل، اندازه و سایر ویژگی‌های آنها در طول زمان تغییر ننماید، ردیابی افراد کار سختی نمی‌باشد [۲]. اما در محیط‌های واقعی، معمولاً شرایط بالا صادق نمی‌باشند. همچنین افراد می‌توانند دچار هم پوشانی شوند و تمام یا بخشی از تصویر آنها برای مدتی توسط تصویر شی دیگر و یا سایر عناصر موجود در محیط پوشانده شود. علاوه بر موارد ذکر شده ردیابی افراد می‌تواند تحت تاثیر تشخیص‌های اشتباه و یا ناقص در الگوریتم تشخیص افراد، منجر به نتایج اشتباه گردد. در خصوص نحوه انتخاب، تولید، ترکیب و چگونگی ارتباط بین ویژگی‌ها جهت تولید یک نتیجه مناسب روش‌های مختلفی ارائه شده است. تفاوت روش‌های ردیابی در مسائلی

^۱ Trajectory^۲ Frame

از قبیل فرضیات صحنه، حالت حرکت شیء و مدل حرکت می باشد که در فصل دوم و سوم به برخی از این روش‌ها اشاره شده است.

عمل ردیابی اهداف متحرک در یک دوربین طی سه مرحله انجام می‌گیرد. در مرحله‌ی اول، اهداف از تصویر استخراج شده و سایر داده‌های اضافی حذف می‌گردند. این مرحله با مشکلاتی از قبیل پویائی محیط، نویز، سایه و هم‌پوشانی بین افراد مواجه می‌باشد. سپس ویژگی‌های نواحی آشکار شده محاسبه می‌گردند. استخراج ویژگی می‌تواند به منظور کاهش حجم اطلاعات و یا متناسب کردن اطلاعات صورت پذیرد. این ویژگی‌ها می‌تواند رنگ، بافت^۱، شکل و یا ویژگی‌های پیچیده‌تری مانند موجک^۲ یا توصیف کننده‌های مستقل از مقیاس^۳ باشند. ثبات ویژگی نسبت به شرایط محیطی، دقت کافی و هزینه‌ی محاسباتی پایین از جمله‌ی مواردی است که در این مرحله حائز اهمیت می‌باشند. در مرحله‌ی سوم ویژگی‌های استخراج شده از یک تصویر در تصویر بعدی تعقیب و دنبال می‌گردند. تکنیک‌های متعددی در این زمینه وجود دارد که در نحوه‌ی جستجوی افراد، مدیریت هم‌پوشانی بین افراد و روش‌های پیش‌بینی با هم تفاوت دارند. در فصل سوم هر یک از این بخش‌ها توضیح داده خواهند شد.

در اغلب موارد سیستم‌های ردیابی، برای نظارت بر محیط‌های بزرگ مانند فرودگاه‌ها، مراکز خرید و محیط‌های هوشمند نیاز است. در این سناریوها نمی‌توان با استفاده از یک دوربین، تمام ناحیه‌ی نظارت را تحت پوشش قرار داد. این محدودیت می‌تواند ناشی از محدودیت‌های وضوح حسگر دریافت کننده‌ی تصاویر (دوربین) و یا موانع موجود در محیط باشد. بنابراین سیستم نظارتی یک ناحیه‌ی بزرگ، باید افراد را در شبکه‌ی ای از دوربین‌ها ردیابی نماید [۲]. همچنین هنگام استفاده از چند دوربین ممکن است پوشش دادن تمامی محدوده‌ی مورد نظر با استفاده از دوربین‌ها میسر نباشد. حتی در

^۱ Texture

^۲ Wavelet

^۳ Scale Invariant Feature Transform (SIFT)

صورتی که بتوان تمام محیط را با استفاده از چندین دوربین با فضای دید مشترک تحت پوشش قرار داد، مشکلاتی نظیر کالیبره^۱ شدن دوربین‌ها با هم به وجود می‌آید که تغییر در فضای دید یک دوربین موجب تغییر در بقیه‌ی دوربین‌ها می‌گردد. بنابر این درساریوهای نظارتی^۲ تعریف شده در محیط‌های واقعی نیاز به سیستم‌هایی است که بتوانند افراد را از طریق چندین دوربین با فضای دید غیر مشترک ردیابی نماید.

ردیابی افراد در دو مرحله انجام می‌پذیرد. ابتدا افراد در تک‌تک دوربین‌ها ردیابی می‌شوند. سپس با استفاده از ویژگی‌های استخراج شده، مشاهدات مربوط به یک فرد مشخص شده و خط سیر افراد در کل محیط محاسبه می‌گردد. به عبارت دیگر در صورتی که مجموعه‌ای از خط سیرها در هر دوربین موجود باشد، باید بتوانیم مشخص کنیم که کدام یک از خط سیرها مربوط به یک فرد خاص می‌باشد.

علاوه بر مسائلی که در ردیابی افراد در یک دوربین مطرح است، ردیابی با استفاده از چندین دوربین از جهات دیگر نیز چالش‌برانگیز است. مشاهدات دریافت شده از یک شی به صورت گسترده‌ای در زمان و مکان قابل توزیع می‌باشند و حضور افراد در زمان و مکان‌های مختلف در ناحیه‌ی دید دوربین‌ها تابع قانون خاصی نمی‌باشد. بنابر این نمی‌توان از ویژگی‌های زمان و مکان برای بررسی شباهت‌ها و مطابقت‌ها استفاده نمود. همچنین، ظاهر یک شی، تابعی از نور محیط، ویژگی‌های هندسی^۳، نوع سطح^۴ شی و مشخصات دوربین می‌باشد. بنابراین ظاهر یک شی در یک دوربین به دلیل عواملی از قبیل شرایط نوری محیط، موقعیت افراد، زاویه‌ی دید و سایر پارامترهای دوربین، می‌تواند با ظاهر همان شی در دوربین دیگر متفاوت باشد.

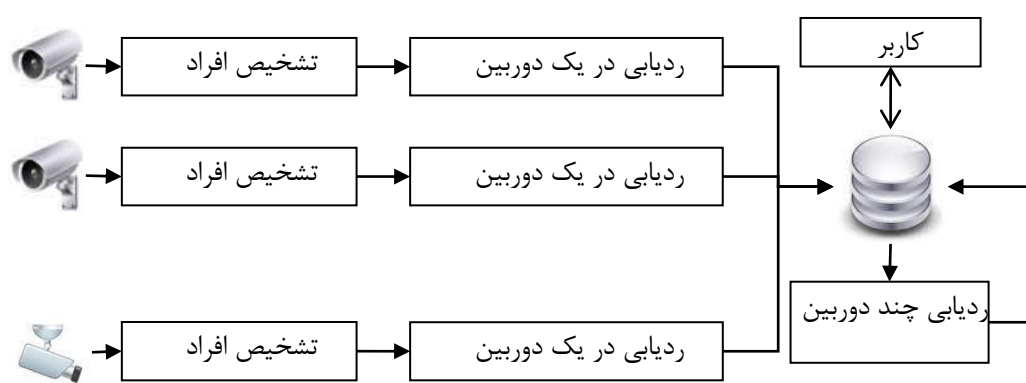
^۱ Calibrate

^۲ Surveillance

^۳ Geometry properties

^۴ Object Surface

در این پایان نامه سیستمی برای ردیابی افراد در شبکه ای از دوربین ها که هیچ اشتراکی در فضای دید ندارند، مورد بررسی قرار گرفته است. ساختار سیستم پیشنهاد شده در شکل ۱-۱ نشان داده شده است. این سیستم از سه زیر سیستم مجزا تشکیل شده است. ردیابی در هر دوربین به صورت مستقل انجام شده و نتایج به دست آمده در قالب یک ساختمان داده ای استاندارد در پایگاه داده ذخیره می شوند. سپس زیر سیستم ردیابی افراد در چند دوربین با استفاده از اطلاعات موجود در پایگاه داده، ارتباط بین مشاهدات و رویدادهای اتفاق افتاده در دوربین های مختلف را برقرار می نماید.



شکل ۱-۱ ساختار سیستم ردیابی پیشنهاد شده و زیر سیستم های آن

۱-۱- اهداف پایان نامه

هدف اصلی این پایان نامه ارائه ی سیستم ردیابی افراد در مجموعه ای از دوربین های تنظیم نشده^۱ با نواحی دید غیرهم پوشان می باشد. سیستم پیشنهاد شده از یک پایگاه داده برای ذخیره ی رویدادها تشکیل شده است. رویدادهای سیستم را می توان به سه دسته ی ورود، خروج و تغییر مکان تقسیم بندی کرد. معمول ترین رویداد، حرکت فرد در محیط دید یک دوربین است. ردیابی افراد در هر یک از دوربین ها به صورت مستقل انجام گرفته و حرکات افراد به همراه جزئیات آن^۲، در قالب یک ساختمان داده ای استاندارد در پایگاه داده ی سیستم ذخیره می گردند. هنگام خروج فرد از محیط دید یک

^۱ Uncalibrated

^۲ جزئیات فرد شامل یک شماره ی منحصر به فرد، اطلاعات زمانی/ مکانی و اطلاعات ظاهری فرد می باشد.

دوربین، ویژگی‌های لازم برای شناسایی در مرحله‌ی بعد شناسایی شده و ذخیره می‌گردند. با ورود یک فرد به محیط دید دوربین، سیستم، ویژگی‌های به دست آمده از فرد را با رکوردهای مربوط به خروج افراد، مقایسه نموده و در صورت تطابق، اطلاعات فرد مورد نظر در پایگاه داده را بروز رسانی می‌نماید. سیستم پیاده‌سازی شده قابلیت دریافت پرس‌وجوهای^۱ مختلف را داشته و خط سیر کلی افراد در مجموعه‌ی دوربین‌ها را بر اساس پارامترهای داده شده تولید می‌نماید. هم چنین این سیستم قابلیت تولید یک فیلم کلی برای یک فرد در مجموعه‌ی دوربین‌ها را دارد.

۲-۱- محدودیت‌ها و فرضیات مساله

در پیاده‌سازی سیستم، شرایط خاصی در انتخاب پایگاه داده در نظر گرفته شده است. اولین شرط، مشخص کننده‌ی نحوه‌ی قرار گیری دوربین‌ها است. دوربین‌ها در مکان‌های ثابتی نصب شده و مکان و زاویه‌ی آنها در حین فیلم برداری تغییر نمی‌نماید. ثابت بودن دوربین این امکان را فراهم می‌آورد که بتوان با استفاده از روش مدل‌سازی پس زمینه، افراد موجود در محیط را تشخیص داد. روشنایی در محیط‌های مختلف متغیر اما تغییرات آن محدود فرض شده است. این تغییرات باید به گونه‌ای باشد که ظاهر شی تغییرات اساسی ننماید. سیستم ارائه شده برای محیط‌های بسته مانند محیط ادارات، موسسات و مراکز خرید در نظر گرفته شده است؛ اما آزمایشات نشان می‌دهد، سیستم قابلیت کار در محیط بیرون را نیز دارد. وجود سایه در این پایان نامه تا حد زیادی حل شده است ولی سایه‌ها نباید خیلی شدید باشند. آشکار سازی هدف، میزان سایه‌ها و نتایج به دست آمده در فصل سوم مورد بررسی قرار گرفته است. شرط دیگر تعداد افراد متحرک در صحنه می‌باشد. تا هنگامی که الگوریتم آشکار سازی افراد در یک دوربین توانایی تشخیص درست افراد را داشته باشد، تعداد افراد بر روی الگوریتم ارائه شده برای برقراری ارتباط در شبکه‌ی دوربین‌ها تاثیری نمی‌گذارد. در حالت کلی تعداد افراد به محدوده‌ی دید دوربین بستگی دارد. الگوریتم ارائه شده در این پایان‌نامه توانایی

^۱ Query

تشخیص و برطرف نمودن هم پوشانی‌های معمول را دارد. اما در صورت پیچیده شدن هم پوشانی بین افراد، سیستم دچار اختلال می‌گردد. حالات ممکن برای هم پوشانی افراد و خروجی سیستم در فصل چهار مورد بررسی قرار گرفته شده است. فرض دیگر این پایان‌نامه ثابت بودن ظاهر شخص در زمان خروج از یک دوربین و زمان ورود به دوربین دیگر تا لحظه‌ی تشخیص است. به این معنی که فرد مجاز به تغییر لباس و نحوه‌ی پوشش خود در محدوده‌ی ورود و خروج محیط نمی‌باشد. پس از این محدوده، فرد به شرطی که سایه‌ی^۱ تشخیص داده شده توسط سیستم تغییر ننماید، می‌تواند نحوه‌ی پوشش خود را تغییر دهد. محدودیتی از لحاظ جهت حرکت افراد وجود ندارد و افراد می‌توانند در جهات مختلف حرکت نمایند. افراد می‌توانند از مسیرهای مختلف وارد سیستم شوند به این شرط که بدن فرد در ناحیه‌ی ورود به صورت کامل در دید دوربین قرار دارد. در مورد زاویه‌ی دوربین‌ها فرض خاصی در نظر گرفته نشده است. اما در صورتی که زاویه‌ی دید دوربین در خروج فرد با ورود فرد تفاوت زیادی داشته باشد، پوشش فرد از روبرو و از پشت سر باید یکنواخت باشد. محدودیت خاصی از لحاظ فاصله‌ی فرد از دوربین در نظر گرفته نشده است. با این وجود اندازه‌ی فرد در صحنه باید از یک مقدار آستانه بزرگتر باشد. در نهایت ذکر این نکته لازم است که هرچند وضوح بالای تصویر باعث بهبود کیفیت ردیابی می‌گردد اما این سیستم با ویدئو‌هایی با ابعاد کوچک (۳۲۰ * ۲۴۰) نیز به خوبی کار می‌نماید.

۱-۳- کاربردها

ردیابی افراد کاربردهای متفاوتی در حوزه‌های مختلف دارد. تمرکز این پایان‌نامه بر روی کاربرد-های ردیابی افراد در سیستم‌های نظارتی و محیط‌های هوشمند می‌باشد. در این بخش برخی از مهمترین کاربردهای ردیابی افراد در حوزه‌های مختلف مورد بررسی قرار گرفته است.

- کنترل بر روی مکان‌های امنیتی

^۱ Silhouette

شاید بتوان ادعا نمود که بیشترین حجم تحقیقات انجام پذیرفته در زمینه ردیابی، کاربرد مراقبتی و امنیتی دارند. تشخیص رفتارهای مشکوک با استفاده از زیر نظر گرفتن حرکات فرد، دنبال کردن افراد در محیط و اعلام خطر هنگام ورود افراد به مناطق پر خطر و زیر نظر گرفتن مجرمین نمونه هایی از کاربرد های نظارتی سیستم های ردیابی می باشند.

- خانه های هوشمند

توانایی تصمیم گیری هوشمند بر اساس رفتار انسان در محیط منزل، یکی از کاربردهای خانه های هوشمند می باشد. ردیابی افراد یکی از اجزاء اصلی این سیستم ها می باشد.

- هوش تجاری ویدئویی^۱

در سالهای اخیر استفاده از سیستم های ویدئویی برای استخراج اطلاعات مفید برای بالا بردن میزان فروش و کسب رضایت مشتریان، توجه بسیاری از محققان و مدیران را به خود جلب نموده است. به عنوان مثال در یک فروشگاه می توان اطلاعاتی از قبیل تعداد افراد در زمان های مختلف در بخش های فروشگاه، میزان زمان صرف شده توسط هر مشتری در هر قسمت، الگوی حرکتی افراد در انتخاب و بازدید از قسمت های مختلف را جمع آوری نموده و با تجزیه و تحلیل آنها اقدامات لازم را در راستای افزایش فروش انجام دهد.

- مراقبت از افراد مسن و ناتوان

مؤسسات حمایت از افراد دارای ناتوانی های جسمی، تلاش گسترده ای را در زمینه ی استفاده از تکنولوژی برای فراهم نمودن زندگی بهتری برای این قشر از جامعه شروع کرده اند. یکی از این تکنولوژی ها استفاده از بینایی ماشین برای مراقبت از این افراد و اتخاذ تصمیمات لازم در مواقع لزوم می باشد.

^۱ Video Business intelligence

۱-۴- ساختار پایان نامه

ادامه‌ی این پایان نامه در پنج فصل گردآوری شده است. در فصل دوم، تکنیک های استفاده شده برای آشکار سازی افراد از پس زمینه مطرح شده و نتایج مربوط به پیاده سازی تکنیک های مناسب برای پایان نامه نمایش داده شده است. فصل سوم، پس از معرفی روش های موجود برای ردیابی افراد در یک دوربین به بررسی نتایج ردیابی افراد در مجموعه داده های تهیه شده می پردازد. سپس در فصل چهارم، روشهای ارتباط اطلاعات به دست آمده از تک تک دوربین ها توضیح داده شده است. این فصل شامل معرفی ویژگی های قابل استفاده در ارتباط اطلاعات و معایب و مزایای آنها می باشد. فصل پنجم نتایج حاصل از سیستم ردیابی ارائه شده را مورد بررسی قرار می دهد. در نهایت در فصل ششم، نتایج حاصل از پایان نامه بررسی شده و پیشنهاداتی برای ادامه ی کار ارائه خواهد شد. به دلیل اینکه پایان نامه در سه فصل مجزا مولفه های سیستم ردیابی افراد را مورد بررسی قرار داده است، معرفی کارهایی که قبلا توسط محققین در مورد هر مولفه انجام پذیرفته، در فصل های مربوطه آورده شده است.

این پایان نامه منجر به مقالات زیر شده است:

1. R.Yousefzadeh, H.Hassanpour., A Generic Model for Smart Surveillance System,. In *The 7th International Workshop on Systems, Signal Processing and their Applications* (Submitted)
2. R.Yousefzadeh, H.Hassanpour., Tracking People through Multiple Cameras with disjoint views, In *3rd International Conference on Machine Learning and Computing* (Accepted)
3. R.Yousefzadeh, H.Hassanpour., Object matching in disjoint cameras using Color Co-Occurrence Matrix, .In *Machine Vision and Applications* (To be submitted)

فصل ۲: آشکارسازی افراد

یک سیستم ردیابی کارآمد، در مرحله‌ی اول نیاز به مکانیزمی برای تشخیص حرکت و آشکار سازی دقیق افراد دارد. در این فصل ابتدا روش‌های موجود در آشکار سازی اهداف بررسی خواهد شد. در بخش دوم مدل ترکیبی گاوسی برای تخمین زمینه و آشکار سازی اشیا که در این پایان نامه پیاده سازی شده، توضیح داده خواهد شد. در نهایت نکاتی در مورد حذف سایه و نویز از تصاویر استخراج شده مورد بررسی قرار خواهد گرفت.

۲-۱- مقدمه

آشکارسازی هدف^۱ از مراحل پایه و اساسی در سیستم های ردیابی می باشد. آشکارسازی هدف در واقع ناحیه‌ی مربوط به هدف مورد نظر را در تصویر قطعه بندی^۲ می‌کند. کارایی مراحل بعدی یک سیستم ردیابی به آشکارسازی هدف بستگی دارند. قطعه‌بندی تصویر یکی از مسائل پیچیده در پردازش تصویر و ویدئو می‌باشد. شرایط دینامیکی محیط مانند تغییرات روشنایی، سایه ها و حرکات متناوب اجسام در اثر وزش باد مانند برگ درخت‌ها آشکارسازی هدف را با مشکل روبرو می‌کند. در این بخش برخی از روش های مطرح در این زمینه معرفی خواهند شد.

یکی از روش‌ها برای آشکارسازی اشیاء، استفاده از اطلاعات یک فریم می‌باشد. در این روش با استفاده از دانش اولیه در مورد خصوصیات اشیاء، پیکسل‌های مرتبط با اشیاء از پیکسل‌های زمینه تمییز داده می‌شوند. در مقابل، روش‌های دیگر از اطلاعات زمانی یا به عبارت دیگر از اطلاعات دنباله-ای از تصاویر برای تشخیص حرکت و آشکارسازی اشیاء از زمینه استفاده می‌نمایند. تحقیقات مربوط به سیستم‌های نظارتی، فرض می‌کنند که افراد موجود در محیط، متحرک هستند. اغلب روش‌هایی که برای آشکارسازی هدف از اطلاعات حرکتی استفاده می‌نمایند، در یکی از سه دسته‌ی زیر تقسیم بندی می‌شوند [۳]:

- شار نوری^۳
- تفاضل زمینه
- تفاضل فریم های متوالی^۴

^۱ Object Detection

^۲ Segmentation

^۳ Optical Flow

^۴ Temporal differencing

الگوریتم‌های موجود برای جداسازی افراد از پس‌زمینه معمولاً دارای خطا می‌باشند؛ بنابراین نیاز به استفاده از پردازش‌هایی برای حذف نویز و اطلاعات ناخواسته مانند سایه می‌باشد. برخی از این روش‌ها در دو بخش پایانی این فصل توضیح داده شده است.

۲-۲- شار نوری

شار نوری به معنی توزیع سرعت جابجایی روشنایی الگوها بین چند تصویر می‌باشد که در آن اجسام متحرک به صورت مجموعه‌ای از بردارهای سرعت نشان داده می‌شوند. در این روش، از بردارهای حرکتی^۱ تغییرات، برای آشکارسازی اهداف استفاده می‌شود. بردارهایی مانند بردار جابجایی می‌توانند به عنوان پایه‌ی الگوریتم‌های ردیابی بر مبنای محیط شیء مورد استفاده قرار بگیرند [۳]. در تصاویر متوالی با ردیابی و جستجوی نقاط متناظر از تصویر اصلی می‌توان جابجایی نقاط متحرک را تحلیل کرد. برای محاسبه‌ی شار نوری نیاز به استفاده از حداقل دو تصویر است. شار نوری یک نقطه از تصویر، بردار جابجایی مکان و سرعت آن نقطه را نتیجه می‌دهد. سپس می‌توان با نگاشت نقطه‌ی مزبور به محیط واقعی، حرکت نقطه‌ی متناظر در محیطی را آنالیز نمود.

شار نوری تخمین نسبتاً دقیقی از افراد متحرک به دست می‌آورد. همچنین از این روش می‌توان برای سیستم‌های با دوربین متحرک نیز استفاده نمود. تخمین بردارهای تصویر بر مبنای شار نوری به شدت نسبت به نویز حساس بوده و دارای محاسبات پیچیده‌ای می‌باشند. در [۴] مفاهیم تئوری شار نوری و روش‌های محاسبه‌ی آن به خوبی مطرح شده است.

۲-۳- تفاضل فریم‌های متوالی

^۱ Flow Vector

این روش از تفاضل دو یا چند فریم متوالی برای آشکار سازی افراد استفاده می‌نماید [۵]. این روش ساده بوده و بار محاسباتی پایینی نسبت به سایر روش‌ها دارد. همچنین در شرایط محیطی پویا به خوبی عمل می‌نماید. با این حال این روش با مشکل روزه‌های شی^۱ مواجه است که پیکسل‌های میانی شی پس از چند فریم به اشتباه به عنوان پیکسل‌های زمینه شناخته می‌شوند [۶]. همچنین تغییرات متوالی پیکسل‌های زمینه مانند حرکت برگ درختان، در این روش به اشتباه تشخیص داده می‌شوند. این روش برای صحنه‌هایی با زمینه‌ی پیچیده مناسب نمی‌باشد.

۲-۴- تفاضل زمینه

در این روش هدف مورد نظر از تفاضل پیکسل‌های تصویر جاری و پیکسل‌های مدل زمینه تعیین می‌شود. این روش کارائی بالایی داشته و برای کاربردهای بی‌درنگ مناسب می‌باشد. در مقابل نسبت به تغییرات پیکسل‌های زمینه حساس می‌باشد. تصویر زمینه ممکن است بنا به دلایلی از جمله تغییر تدریجی روشنایی محیط، تغییر سایه‌ی اشیاء به دلیل تغییر زاویه‌ی خورشید، تغییرات ناگهانی در آب و هوا مانند عبور ابر از مقابل خورشید، و تغییرات غیر طبیعی مانند روشن و یا خاموش شدن چراغ‌ها تغییر نماید. یکی دیگر از معایب این روش حساسیت زیاد نسبت به مقدار آستانه می‌باشد. در سال‌های اخیر روش‌های بسیاری، برای بهبود و رفع معایب این روش مطرح شده‌اند. در این روش‌ها سعی می‌شود تخمین مناسبی از زمینه در هر فریم بدست آورده شود. روش‌هایی از جمله میانگین گاوسی مداوم [۷]^۲، فیلتر زمانی میانه^۳، مقادیر ویژه‌ی زمینه^۴ [۸] برای حل مشکلات روش تفاضل زمینه ارائه شدند. در این میان روش ترکیبی گاوسی دارای دقت و کارائی بالایی در مدل‌سازی زمینه دارد. به همین دلیل در این پایان‌نامه از این روش استفاده شده است. این روش می‌تواند بر مشکلاتی از قبیل

^۱ Foreground aperture

^۲ Running Gaussian Average

^۳ Temporal Medium Filter

^۴ Eigen Background

تغییرات متناوب نور که توسط لامپ‌های فلورسنت و حرکات کوچک متناوب مانند حرکت برگ درختان، غلبه نماید. همچنین این روش می‌تواند اشیائی را که برای مدت طولانی در محیط باقی می‌مانند، مدیریت نماید. در این روش مقدار هر پیکسل، با مجموعه‌ای از چند توزیع گاوسی مدل می‌گردد [۹]. این توزیع‌ها در هر فریم بروز رسانی می‌شوند. در ادامه‌ی فصل، این روش به تفصیل توضیح داده شده است.

روش‌های دیگری نیز وجود دارند که از چندین سطح یا چندین پارامتر برای آشکار سازی تغییرات استفاده می‌نمایند. به عنوان مثال الگوریتم Walflower در سه سطح تغییرات زمینه را مدیریت می‌نماید. این سه سطح عبارتند از تغییرات در سطح پیکسل‌ها، تغییرات ناحیه‌ای و تغییرات در کل تصویر. سیستم W4، با استفاده از سه پارامتر حداقل شدت، حداکثر شدت و حداکثر شدت تغییرات بین فریم‌های متوالی، زمینه را به صورت آماری مدل سازی می‌نماید [۱۰]. این سیستم برای تشخیص و ردیابی افراد در محیط‌های بیرون که تغییرات کمی در شرایط نوری و زمینه وجود دارد، طراحی شده است.

در این بخش توضیح مختصری راجع به متداولترین روش‌های مدل سازی زمینه ارائه شد. روش‌های دیگری نیز وجود دارند که با توجه به شرایط محیط و کاربرد‌های متفاوتی طراحی شده‌اند. در [۱۱] و [۱۲] برخی از روش‌های موجود مقایسه شده‌اند.

۲-۵- تخمین پس زمینه به کمک مدل ترکیبی گاوسی^۱

مدل ترکیبی گاوسی یک روش کارآمد برای مدل کردن سیگنال و تصویر است که اولین بار توسط Stauffer و Grimson برای مدل‌سازی زمینه مورد استفاده قرار گرفته است [۹]. این مدل مانند یک مدل مخفی مارکف (HMM) یک حالت است که تابع چگالی احتمال آن دارای چند توزیع نرمال می‌-

^۱ Mixture of Gaussian

باشد. احتمال تعلق بردار X به یک مدل ترکیبی گاوسی دارای M توزیع نرمال به شکل زیر بیان می شود:

$$P(X | GMM) = \sum_{i=1}^M c_i \cdot N(\mu_i, \Sigma_i) \quad (1-2)$$

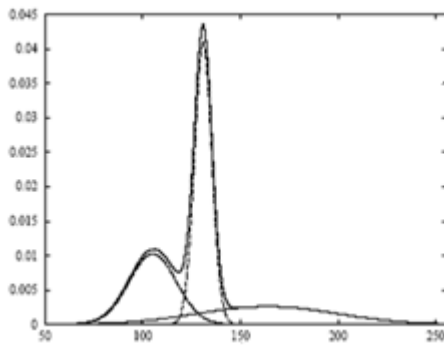
که در آن c_i وزن مدل گاوسی i و μ_i, Σ_i به ترتیب بردار میانگین و ماتریس کواریانس توزیع نرمال هستند. کواریانس مدل ترکیبی گاوسی معمولاً به صورت قطری در نظر گرفته می شود. گرچه امکان استفاده از ماتریس کامل نیز وجود دارد. رابطه $(1-2)$ را می توان با استفاده از تابع چگالی احتمال نرمال به صورت زیر بیان نمود:

$$P(X | GMM) = \sum_{i=1}^M c_i \cdot \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_i|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(x_i - \mu_i)^T \Sigma_i^{-1} (x_i - \mu_i)\right) \quad (2-2)$$

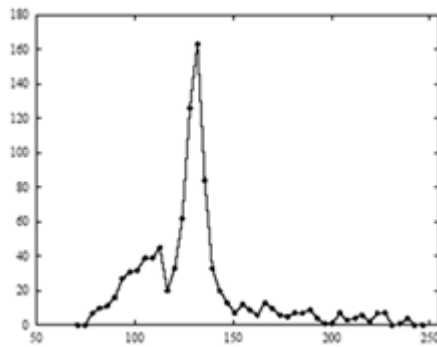
که در آن d ، بعد فضای ورودی است. برای به دست آوردن پارامترهای مدل ترکیبی گاوسی شامل وزن مدل گاوسی و میانگین و کواریانس توزیع ها از الگوریتم ماکزیمم نمودن امید ریاضی^۱ (EM) استفاده شده است.

شکل ۱-۲ (الف) چگالی مربوط به یک پیکسل در دویست و پنجاه فریم متوالی را به تصویر کشیده است. شکل ۱-۲ (ب) مدل گاوسی تخمین زده شده برای مقادیر نشان داده شده در شکل الف را نشان می دهد. این تخمین شامل سه توزیع گاوسی است. توزیع گاوسی که بزرگترین میانگین را دارد رنگی را مشخص می کند که در فریم های متوالی بیشتر از همه تکرار شده است و مربوط به پس زمینه است. به همین ترتیب توزیع گاوسی بعدی می تواند مربوط به سایه و یا مقدار تولید شده از جابجایی قسمت های متحرک (مانند برگ ها) باشد.

^۱ Expectation Maximization



(ب)



(الف)

شکل ۱-۲ (الف) چگالی مربوط به یک پیکسل در دویست و پنجاه فریم متوالی (ب) توابع های گاوسی تشخیص داده شده برای مقادیر موجود [۹]

۲-۵-۱- تئوری مدل گاوسی^۱

اگر دنباله‌ی شدت رنگ یک پیکسل در طول زمان، $x_1, x_2, \dots, x_n$ باشد، تابع چگالی احتمال برای x_t در زمان t ، با استفاده از تابع تخمین K به صورت زیر محاسبه می شود:

$$p(x_t) = \frac{1}{n} \sum_{i=1}^n k(x_t - x_i) \quad (۳-۲)$$

احتمال این که پیکسلی شدت رنگ x_t را در زمان t داشته باشد، با توزیع گاوسی به صورت زیر تعریف می شود:

$$p(x_t) = \sum_{i=1}^k (w_{i,t} * f(x_t) \mu_{i,t}, \sum_{i,t}) \quad (۴-۲)$$

که در آن K تعداد توزیع، $w_{i,t}$ وزن تخمین زده شده برای آامین توزیع، $\mu_{i,t}$ میانگین آامین توزیع در زمان t و $\sum_{i,t}$ ماتریس کواریانس آامین توزیع در زمان t و بالاخره f تابع توزیع گاوسی به صورت زیر می باشد:

^۱ Mixture of Gaussian model

$$f(x_t | \mu_k, \sum_k) = \frac{1}{(2\pi)^{\frac{n}{2}} \left| \sum_k \right|^{\frac{1}{2}}} e^{-\frac{1}{2} (x_t - \mu_k)^t \sum_k^{-1} (x_t - \mu_k)} \quad (5-2)$$

که n برابر بعد بردار x می باشد. برای کم شدن بار محاسباتی موجود در این روش فرض می شود اجزاء بردار x (نمونه های پیکسل) که همان رنگ های، قرمز و سبز و آبی هستند دارای واریانس یکسان و مستقل از هم هستند. بنابراین ماتریس کواریانس $(\sum_{i,t})$ به صورت زیر تعریف می شود:

$$\sum_k = \delta_k I \rightarrow \sum_k = \begin{pmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \sigma_g^2 & 0 \\ 0 & 0 & \sigma_b^2 \end{pmatrix} \quad (6-2)$$

مزیت این روش این است که برای نگهداری چگالی شدت رنگ یک پیکسل نیاز به نگهداری تمام پیکسل ها نبوده و با متغیرهای میانه و میانگین می توان تابع چگالی آن را تشکیل داد.

۲-۵-۲- به روزرسانی مدل ترکیبی گاوسی^۱

δ_i, μ_i, W_i پارامترهای i امین تابع گاوسی می باشند که در هر مرحله باید بروزرسانی شوند. در صورتی که مقدار پیکسل X_t با k توزیع مربوط به آن مقایسه و مقدار پیکسل در معادله (۷.۲) صدق کرد، پارامترهای ذکر شده باید بروز رسانی گردند.

$$\frac{X_t - \mu_t}{\delta_t} > 2.5 \quad (7-2)$$

مقدار $2/5$ یک حد آستانه‌ی تجربی است که توسط نویسندگان مقاله پیشنهاد شده است. در این هرگاه مقدار جدیدی مانند نویز در پس زمینه ظاهر شود، باعث خرابی و از بین رفتن زمینه نمی شود.

^۱ Mixture model updating

اگر هیچ یک از K توزیع با مقدار پیکسل جدید (X_t) تطبیق نداشته باشند، مقدار جدید جایگزین کوچکترین احتمال توزیع می گردد.

در هر مرحله هر یک از وزن ها از وزن قبلی تأثیر می پذیرند. همچنین توزیعی که با مقدار پیکسل جدید تطبیق داده شده است به اندازه α برابر بیشتر از وزن های دیگر تحت تأثیر قرار می گیرد.

$$W_{k,t} = (1 - \alpha)W_{k,t-1} + \alpha m(w_k)X_t \quad (۸-۲)$$

در رابطه ی (۸.۲) α نرخ یادگیری و m در صورت تطبیق مقدار پیکسل با انحراف معیار، مقدار ۱ و در غیر این صورت مقدار صفر بر می گرداند:

$$m(w_k|X_t) = \begin{cases} 1 & \text{if } X_t \text{ matches distribution } k \\ 0 & \text{otherwise} \end{cases} \quad (۹-۲)$$

مقادیر میانگین و انحراف معیار نیز طبق فرمول های زیر بروز رسانی می شوند :

$$\mu_{k,t} = (1 - f_k)\mu_{k,t-1} + f_k X_t \quad (۱۰-۲)$$

$$\delta_{k,t}^2 = (1 - f_k)\delta_{k,t-1}^2 + \eta_k (X_t - \mu_{k,t})^t (X_t, \mu_{k,t}) \quad (۱۱-۲)$$

۲-۵-۳- آشکار سازی پیش زمینه^۱

برای آشکار سازی افراد، پس از مرتب سازی k توزیع به صورت $\frac{w_i}{\delta_i^2}$. b توزیع اول به عنوان مدل

زمینه مورد استفاده قرار می گیرد:

$$B = \arg \min \left(\sum_{k=1}^b w_k \right) T \quad (۱۲-۲)$$

^۱ Foreground Segmentation

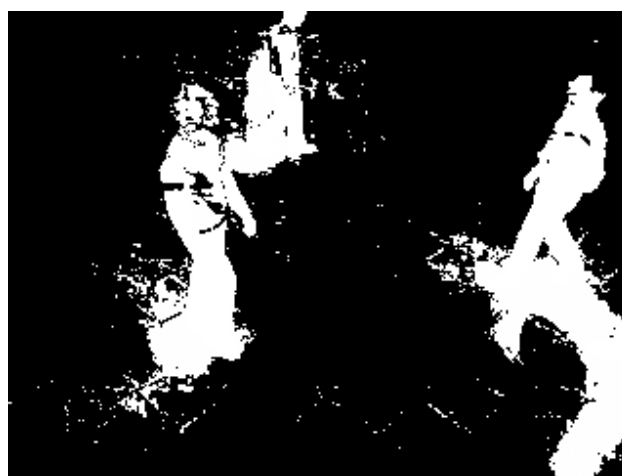
T در رابطه‌ی (۱۲.۲) حد آستانه می باشد. به این معنی که مقادیر وزن k توزیع را تا جایی به هم اضافه می‌نماییم که به مقدار آستانه برسیم. توزیع B برای تمامی پیکسل‌ها محاسبه می‌شود. هر پیکسلی که مقدار آن با مقدار B منطبق بود، به عنوان زمینه در نظر گرفته می‌شود. در شکل ۴-۲ نمونه‌ای از آشکارسازی شیء از زمینه آورده شده است. برای محاسبه‌ی تفاضل، توزیع‌های تولید شده از دنباله‌ی فریم‌ها با شروع از شکل ۳-۲ به عنوان زمینه و شکل ۲-۲ به عنوان فریم فعلی در نظر گرفته شده است.



شکل ۳-۲ تصویر زمینه، مدل زمینه با این تصویر برابر نبوده و مجموعه‌ای از مدل‌های گاوسی برای هر پیکسل می باشد



شکل ۲-۲ فریم فعلی



شکل ۴-۲ نتیجه‌ی تفاضل فریم فعلی و مدل زمینه

۲-۶- حذف نویز

به دلیلی وجود فاکتورهای مختلف در مدل سازی زمینه، همواره مقداری نویز در نتایج وجود دارد. متداول ترین روشی که سیستم های ردیابی برای حذف نویز از آن بهره می گیرند، فیلتر های ریخت شناسی^۱ فرسایش^۲ و گسترش^۳ و یا ترکیبی از این دو است. این دو فیلتر بر روی تصویر به دست آمده از مراحل قبل که یک تصویر دودویی شامل پس زمینه و اشیا می باشد، اعمال می شوند. این فیلتر ها بر روی مرز های اشیا تغییراتی را اعمال می نمایند. فیلتر فرسایش برای از بین بردن نویز های ضربه ای کاربرد دارد. فیلتر گسترش نیز می تواند اجزاء متعلق به یک شی را که به دلیل اشتباه در قطعه بندی، به دو قسمت مجزا ی نزدیک به هم تقسیم شده اند، به هم متصل نماید.

۲-۷- حذف سایه

در اغلب موارد، سایه ی فرد نیز به عنوان قسمتی از فرد شناخته می شود؛ در حالی که سایه قسمتی از زمینه ی تصویر می باشد. بنابراین حذف سایه تاثیر بسزایی در افزایش کارایی سیستم ردیابی خواهد داشت. الگوریتم های متنوعی برای حذف سایه وجود دارد. به عنوان مثال در [۱۳] فرض شده است هنگامی که سایه ای بر روی یک شی می افتد، شدت سطح خاکستری آن شی تغییر می نماید، در حالی که مقدار رنگ نرمال شده تغییر قابل توجهی نمی کند. به عبارت دیگر اگر $I(x,y)$ مقدار شدت سطح خاکستری را در نقطه (x,y) نشان دهد، خواهیم داشت :

$$\frac{R_s}{R_s + G_s + B_s} \cong \frac{R}{R + G + B}, \frac{G_s}{R_s + G_s + B_s} \cong \frac{G}{R + G + B}, \frac{B_s}{R_s + G_s + B_s} \cong \frac{B}{R + G + B} \quad (۱۳-۲)$$

^۱ morphologic^۲ Erosion^۳ Dilation

$$I_s(x, y) = \alpha I(x, y) \quad (14-2)$$

در روش دیگری که در مقاله‌ی [۱۴] معرفی شده است، میزان درخشندگی هر پیکسل به عنوان معیاری برای تشخیص سایه استفاده شده است. هنگام افتادن سایه بر روی یک سطح، مولفه‌ی درخشندگی بیشتر از مولفه‌هایی رنگی نور تحت تاثیر قرار می‌گیرند. اگر تصویر را به فضای رنگ HSV نگاشت نماییم، با استفاده از رابطه زیر می‌توان سایه را از تصویر جدا کرد.

$$shadow_{x,y}^k = \begin{cases} 1 & \text{if } D_H < \tau_h, |I_{x,y_s}^k - B_{x,y_s}^k| < \tau_s, v_a \leq \frac{I_{x,y_v}^k}{B_{x,y_v}^k} \leq v_b \\ 0 & \text{otherwise} \end{cases} \quad (15-2)$$

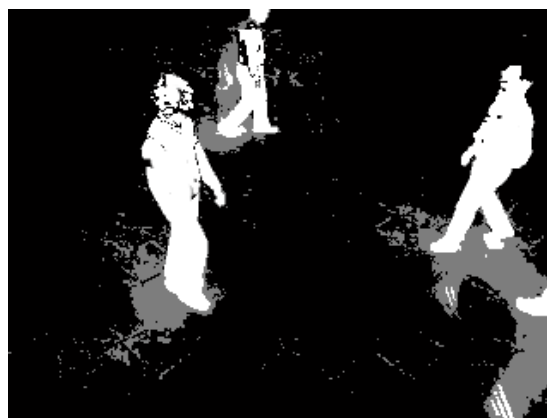
$$D_H = \min \left\{ |I_{x,y-h}^k - B_{x,y-h}^k|, 360 - |I_{x,y-h}^k - B_{x,y-h}^k| \right\} \quad (16-2)$$

در عبارات فوق $I_{x,y}^k$ پیکسل واقع در موقعیت (x,y) فریم k ام و $B_{x,y}^k$ پیکسل واقع در موقعیت (x,y) پس زمینه را در فریم k ام نشان می‌دهد. مقادیر τ_h, τ_s, v_a, v_b اعدادی در بازه‌ی صفر و یک می‌باشند که با استفاده از آن می‌توان کران مقادیر را در رابطه (۱۵.۲) کنترل کرد.

شکل ۵-۲ نتیجه‌ی حاصل از تشخیص سایه توسط الگوریتم اول را نشان می‌دهد. در این تصاویر قسمت خاکستری ناحیه‌ای از خروجی مرحله‌ی تشخیص اشیا است که به عنوان سایه شناخته شده است. همانگونه که مشاهده می‌نمایید حجم عمده‌ای از پیکسل‌ها، به عنوان سایه تشخیص داده شده‌اند.



(الف)



(ب)



(پ)

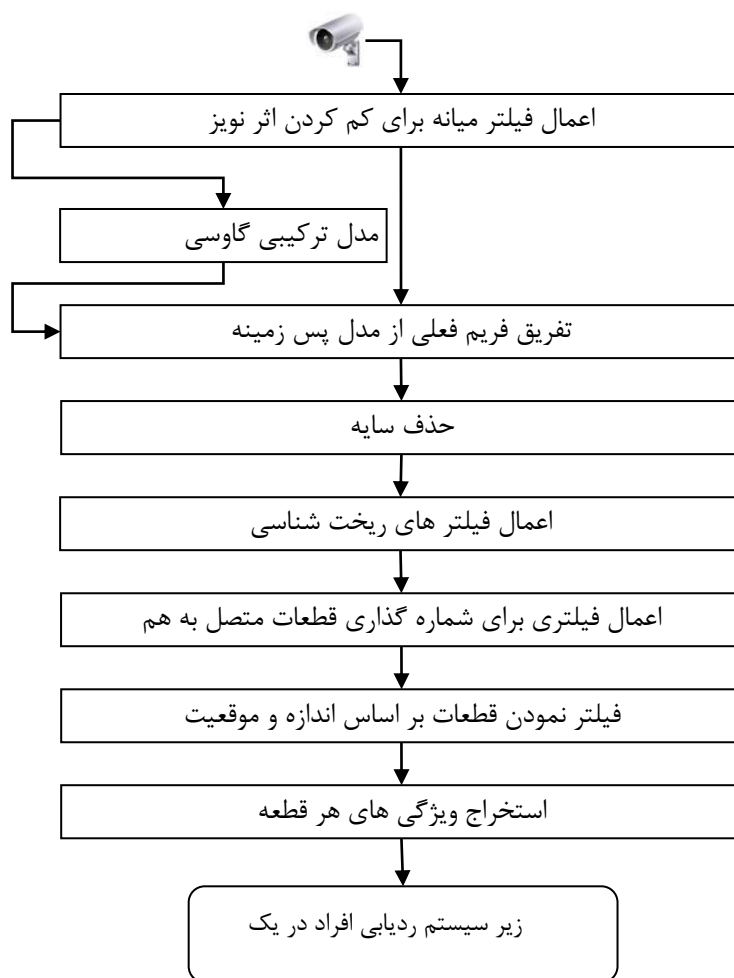


(ت)

شکل ۲-۵ دو فریم و سایه ی تشخیص داده شده، قسمت های خاکستری نشان دهنده ی سایه می باشند. بخش الف و پ تصاویر اصلی و بخش ب و ت سایه های تشخیص داده شده هستند.

۲-۸- شمای کلی زیر سیستم آشکار سازی افراد

شکل زیر شمای کلی زیر سیستم آشکار سازی افراد را که در این پروژه پیاده سازی شده است، نمایش می دهد. پس از آشکار سازی افراد در محیط به وسیله ی تفاضل فریم فعلی از مدل زمینه که با استفاده از مدل ترکیبی گاوسی، پردازش هایی برای حذف نویز و سایه انجام می پذیرد. در مرحله ی بعد، ابتدا مولفه های متصل شماره گذاری شده و سپس بر اساس اندازه فیلتر می شوند. اگر اندازه ی آنها از حد آستانه بزرگتر بودند، ویژگی های آنها استخراج شده و به زیر سیستم ردیابی در یک دوربین ارسال می گردند.



شکل ۲-۶: زیرسیستم آشکار سازی اهداف

۲-۹- نتیجه گیری

در این فصل، زیر سیستم آشکار سازی افراد که اولین مرحله از یک سیستم ردیابی است مورد بررسی قرار گرفت. ابتدا برخی از روش های معمول از جمله شار نوری و تفاضل فریم ها برای آشکار سازی افراد توضیح داده شده و مزایا و معایب هر یک ذکر گردید. سپس مدل ترکیبی گاوسی به عنوان یکی از روش های موفق برای مدل سازی زمینه مورد بررسی قرار گرفته و نحوه ی آشکار سازی اهداف با آن توضیح داده شد. جهت بالا بردن کارایی، پردازش هایی برای حذف نویز و سایه معرفی شده و در نهایت روال کلی زیرسیستم توضیح داده شد. استفاده از روش ترکیب گاوسی به همراه پردازش های معرفی شده می تواند علاوه بر نیاز های یک محیط بسته، اشیا را در محیط های باز نیز آشکار نماید.

فصل ۳: ردیابی افراد در یک دوربین

ردیابی افراد به معنی استخراج ویژگی های یک شی در یک فریم و تعقیب شی در فریم های بعدی بر اساس ویژگی های استخراج شده می باشد. در این فصل مبانی ردیابی اهداف در یک دوربین مورد بررسی قرار خواهند گرفت. در این فصل پس از مروری بر روش های موجود برای ردیابی اهداف و بررسی مزایا و معایب هر روش، فیلتر کالمن به عنوان راهکاری برای ردیابی افراد در حضور نویز و عوامل مخربی چون هم پوشانی افراد مورد بحث قرار خواهد گرفت. در انتهای فصل نیز راهکارهایی برای ردیابی اهداف چند گانه مطرح شده است

۳-۱- مقدمه

پس از اینکه افراد از زمینه‌ی تصویر جدا شدند، می‌توان آنها را در محیط تعقیب نموده و خط سیر آنها را به دست آورد. روش‌های بسیار زیادی برای ردیابی اهداف در کاربردهای مختلف ارائه شده‌اند. تمرکز این فصل بر روی ردیابی افراد از طریق ارتباط احتمالی داده‌ها^۱ و فیلتر کالمن می‌باشد. روش‌های ردیابی افراد را می‌توان به دو دسته‌ی روش‌های قطعی و روش‌های احتمالی تقسیم نمود. در روش‌های قطعی از برخی روابط برای برقراری ارتباط بین اشیاء در دو فریم متوالی استفاده می‌نمایند. اما داده‌های دریافت شده از سنسورها، کاملاً دقیق نبوده و دارای نویز می‌باشند که سیستم ردیابی را دستخوش تغییرات ناخواسته می‌نمایند. روش‌های آماری، علاوه بر داده‌های اندازه‌گیری شده، عدم قطعیت را نیز در روال تشخیص و پیش‌بینی اشیاء مشابه در فریم‌های متوالی مد نظر قرار می‌دهند. در این روش‌ها عموماً از یک فضای حالت برای مدل کردن ویژگی‌های شیء مانند مکان، سرعت و شتاب استفاده می‌نمایند. حالات اشیاء در تصویر مانند ظاهر، موقعیت، سرعت و یا سایر مشخصات بر اساس حالات تشخیص شده قبلی و معادلات تغییر حالات تخمین زده می‌شود. تخمین به دست آمده با حالت اشیاء یافته‌شده در تصویر واقعی، مقایسه شده و فضای حالت بروز رسانی می‌گردد. در این فصل، پس از مرور مختصری بر روش‌های موجود، در مورد نحوه‌ی استفاده از مدل‌های آماری برای پیش‌بینی ویژگی‌های مربوط به یک شیء بحث خواهد شد.

۳-۲- مروری بر کارهای گذشته

مشکلات عمده‌ی ردیابی افراد را می‌توان در شش دسته‌ی زیر طبقه‌بندی نمود[۳]:

۱. تعداد افراد تشخیص داده شده
۲. جدا سازی و ردیابی افرادی که به صورت گروهی حرکت می‌نمایند.

^۱ Probabilistic data association

۳. هم پوشانی جزئی افراد، یعنی یک فرد یا قسمتی از صحنه، بخشی از سایه ی فرد را بپوشاند.

۴. هم پوشانی کلی افراد، هنگامی که فرد کاملاً توسط یک شی دیگر پوشانده شود.

۵. دقت پایین روش قطعه‌بندی افراد که منجر به تشخیص‌های ناقص و مخدوش از افراد می‌گردد.

دلایل متعددی از جمله کنتراست پایین بین ظاهر فرد و صحنه، نویز و سایه می‌توانند موجب این مشکل گردند.

۶. شباهت متقابل بین افراد موجود در سیستم که می‌تواند باعث ارتباط داده‌ی اشتباه گردد.

مقالات و کارهای انجام شده در زمینه‌ی ردیابی افراد در یک دوربین را می‌توان از دو جهت مورد بررسی قرار داد: نحوه‌ی انتخاب معیار ردیابی و نحوه‌ی تطابق مشاهدات. در ادامه به بررسی هر مورد پرداخته خواهد شد.

۳-۲-۱- نحوه‌ی انتخاب معیار ردیابی

از دید نحوه‌ی انتخاب معیار ردیابی، روش‌های ردیابی را می‌توان به دو دسته‌ی ردیابی بر اساس ظاهر و ردیابی بر اساس مدل تقسیم بندی نمود.

الف) ردیابی بر اساس ظاهر

در این روش‌ها ابتدا ویژگی‌های ظاهری هدف در یک فریم استخراج شده و از آن برای تطابق بین مشاهدات در فریم‌های بعدی استفاده می‌شود [۳]. در این روش مسائلی از قبیل انتخاب ویژگی، چگونگی تطابق بین ویژگی‌ها و نحوه‌ی بروز رسانی ویژگی‌ها مطرح می‌باشد. این روش را می‌توان بر اساس انتخاب ویژگی‌ها به سه گروه زیر طبقه بندی کرد [۳]:

- ردیابی بر مبنای ویژگی‌های کلی^۱:

^۱ Global Feature-Based Algorithm

در این روش ویژگی‌هایی مانند، مرکز هدف [۱۵]، پیرامون، مساحت و رنگ [۱۶]، [۱۷]، [۱۸]، مورد استفاده قرار می‌گیرد. در مقاله‌ی [۱۵] هر فرد توسط یک مستطیل^۱ احاطه شده و مرکز آن به عنوان ویژگی برای ردیابی استفاده می‌شود. هنگامی که برخوردی بین دو فرد رخ می‌دهد با توجه به سرعت، نقطه مرکزی هر فرد تشخیص داده می‌شود و ردیابی به درستی انجام می‌شود.

- ردیابی بر مبنای ویژگی‌های محلی^۲ که از ویژگی‌هایی مانند نقاط، خطوط و گوشه‌ها مورد استفاده قرار می‌گیرند.

- ردیابی بر مبنای وابستگی گراف^۳ که شامل مجموعه‌ای از فاصله بوده و بر روی ارتباط ژئومتریک ویژگی‌ها تمرکز دارد

برای بهبود کارایی می‌توان این سه روش را ترکیب نمود. همچنین مدل مخفی مارکف^۴، شبکه‌های عصبی مصنوعی^۵ (ANN)، پارسیکل فیلتر، فیلتر کالمن و الگوریتم جابجایی میانگین^۶ الگوریتم‌هایی هستند که برای ردیابی ویژگی‌ها استفاده می‌شوند. روشهای مبتنی بر ویژگی به دلیل تکیه بر ظاهر افراد، در صورت وجود همپوشانی جزئی می‌توانند افراد را به درستی تشخیص دهند.

در برخی از روش‌های ارائه شده، دوربین می‌تواند متحرک باشد. در این روش‌ها باید ویژگی‌هایی انتخاب شوند که بتوان بدون استفاده از تفاضل فریم آنها را استخراج نمود. به عنوان مثال در [۱۹] روشی ارائه شده است که دوربین می‌تواند نسبت به هدف آزادانه حرکت نماید. در [۲۰] از مدل مخفی مارکف برای مدل کردن مرزهای شکل استفاده شده و الگوریتم ردیابی JPDAF پارامترهای خود را از این مدل می‌گیرد.

^۱ Rectangular Box

^۲ Local Feature-Based Algorithm

^۳ Dependence-Graph-Based Algorithm

^۴ Hidden Markov model

^۵ Artificial Neural Network

^۶ Mean Shift

به عنوان مثالی از ردیابی به وسیله‌ی گراف وابستگی می‌توان به مقاله‌ی [۲۱] اشاره نمود. در این مقاله مسأله‌ی ردیابی، بصورت یک مسئله برچسب دهی گراف‌ها بر روی یک گراف نواحی مجاور^۱ (RAG)، تولید شده بر مبنای اطلاعات حرکت مدل شده است.

در [۲۲] یک سری از ویژگی‌ها مانند شکل، رنگ و ویژگی‌های لبه برای هدف در نظر گرفته شده و فیلتر کالمن برای ردیابی اشیاء استفاده نموده‌اند. آشکارسازی و ردیابی با استفاده از رنگ یکی از روش‌های مهم در کاربردهای ردیابی به صورت بلادرنگ می‌باشد. ردیابی با استفاده از رنگ دارای مزایای بسیاری نظیر بار محاسباتی کم و استقلال نسبت به چرخش و تغییر شکل شیء متحرک است. با این وجود تحت شرایط محیطی متغیر، ردیابی با این روش مشکل می‌باشد. برای غلبه بر این مشکل روش‌های مختلفی پیشنهاد شده است که همه از ترکیب ایده‌های گوناگون با روش مبتنی بر رنگ ایجاد می‌شوند [۱۸].

موجک^۲ نیز یکی دیگر از ویژگی‌هایی است که از آن به عنوان یک ویژگی با ثبات نسبتاً بالا در کاربردهای ردیابی افراد از آن استفاده شده است. نمونه‌هایی از این ویژگی در مقالات [۲۳] و [۲۴] مورد بررسی قرار گرفته‌اند.

گروهی از روش‌ها ناحیه‌ی هدف تصویر را با استفاده از اطلاعاتی از جمله مدل حرکت، رنگ و بافت، تحت ردیابی قرار می‌گیرد. معمولاً هدف مورد نظر از طریق تفریق دو تصویر آشکارسازی می‌شود. نواحی تصویر را می‌توان به شکل‌های مختلفی در نظر گرفت. Wren، از حباب‌های کوچک برای ردیابی یک فرد در تصویر استفاده کرده است [۲۵]. در این مقاله بدن انسان با حباب‌های کوچکی که هر یک معرف قسمتی از بدن انسان است نمایش داده شده است. بنابراین با ردیابی هر یک از حباب‌ها در واقع هدف مورد نظر ردیابی می‌شود. مقدار هر پیکسل توسط یک تابع گاوسی

^۱ Region Adjacent Graph

^۲ Wavelet

مدل شده است. در این مقاله هیچ راهکاری برای غلبه بر هم پوشانی بین افراد در نظر گرفته نشده است. بنابراین افراد باید بدون تماس با یکدیگر در حال حرکت باشند.

در مقاله‌ی [۲۶] حرکت هر شیء دو بعدی در فضا را با یک تغییر شکل دو بعدی، انتقال و چرخش دو بعدی مدل شده و از فاصله هاسدورف^۱ برای تطبیق ناحیه استفاده کردند. الگوریتم فاصله هاسدورف یک الگوریتم برای تطابق اشکال است که سعی می کند از بین همه راه های ممکن انتقال و چرخش، کوتاهترین راهی که یک مجموعه از نقاط به مجموعه نقاط دیگری تصویر می کند را بیابد. این روش در مقابل تغییر شکل و چرخش مقاوم بوده و در مقابل بار محاسباتی بالایی دارد.

در روشی دیگر از اطلاعات رنگ و گرادیان برای آشکارسازی استفاده شده و ردیابی در سه سطح ناحیه، مردم و گروه انجام شده است [۲۷]. یک فرد با چند ناحیه که توسط یک مستطیل محاط^۲ مشخص شده اند، به هم پیوسته می شود. ردیابی در این روش بر مبنای رنگ هر ناحیه می باشد.

پیرامون اشیا^۳ نیز یکی از ویژگی هایی است که مورد توجه بسیاری از محققین قرار گرفته است. در این گروه ردیابی هدف به وسیله استخراج پیرامون هدف و بروز رسانی دینامیکی آن در فریم های بعدی می باشد. در این روش جداسازی اشیایی که به طور جزئی همپوشانی دارند با مشکل روبرو خواهد شد اما اگر در هنگام همپوشانی بتوان برای هر شیء یک کانتور جدید ایجاد نمود می توان اشیاء را ردیابی نمود.

پس از استخراج پیرامون شی می توان از ترکیب روش های دیگر برای ردیابی افراد استفاده نمود. به عنوان مثال در [۲۸] از فیلتر کالمن برای ردیابی هدف های غیر صلب^۴ مثل انسان استفاده شده است. به عنوان مثالی دیگر مقاله ی [۲۹]، یک چهارچوب ردیابی برای اشیاء صلب و غیر صلب در

^۱ Hausdorff Distance Algorithm

^۲ Bounding Box

^۳ Contour

^۴ Non Rigid

حضور هم پوشانی، ارائه داده است که کانتور شیء به زیر کانتورهایی تقسیم می شود و با ترکیب مدل کانتور فعال و وزن دهی به خصیصه های سطح پایین رنگ و حرکت، اشیا را در محیط تعقیب می نماید.

ب) ردیابی مبتنی بر مدل

در روش های مبتنی بر مدل، یک مدل دو بعدی یا سه بعدی از شیء ردیابی می گردد. مدل شیء معمولاً طی یک فرایند تکراری با تصویر تطبیق داده شده و بهترین محل تطابق آن به عنوان محل مشاهده شیء در نظر گرفته می شود. این دسته از روش ها بیشتر برای ردیابی وسائط نقلیه و اعضای بدن انسان [۱۰] به کار رفته اند. سیستم های ردیابی که از مدل استفاده نمی نمایند، براساس روش های تشخیص حرکت نواحی متحرک را از تصویر جدا نموده و بر اساس ویژگی های استخراج شده از ناحیه ی مربوط به هر شیء، عمل ردیابی انجام می پذیرد. این سیستم ها به سطح فهم تصویر نمی رسند و تنها گروهی از پیکسل ها را ردیابی می نمایند. در مقابل سیستم های مبتنی بر مدل به سطحی از فهم تصویر دست می یابند. مدل اهداف معمولاً از طریق اندازه گیری های دستی^۱، نرم افزارهای CAD^۲ و یا تکنیک های بینایی ماشین^۳ ایجاد می شود.

اگر مدل مناسبی بر اساس مشاهدات و دانش اولیه برای هدف ایجاد شود، ردیابی دقیق انجام خواهد شد. همچنین استفاده از مدل، امکان تشخیص و رفع هم پوشانی افراد را نیز فراهم می نماید. در این روش اطلاعات مربوط به ساختار بدن انسان، محدودیت های حرکتی و اطلاعات اولیه می تواند با هم ترکیب شده و الگوریتم ردیابی را تقویت نمایند. در مقابل ایجاد یک مدل برای هدف و جستجوی آن در تصویر نیازمند حجم محاسباتی بالایی می باشد.

^۱ Manual measurement

^۲ Computer Aided Design Tools

^۳ Computer vision techniques

ردیابی مبتنی بر مدل برای اهداف صلب و غیرصلب تفاوت هایی دارد. ردیابی مبتنی بر مدل اهداف غیر صلب مانند بدن انسان^۱ مبتنی بر مدل آنالیز به وسیله پیوند^۲ می باشد که برای ردیابی از شیوه-ی تخمین- تطبیق- بروزرسانی^۳ استفاده می کند. ابتدا موقعیت مدل در فریم فعلی توسط دانش اولیه از هدف تخمین زده می شود. با استفاده از روش های جستجو^۴، مدل تخمین زده شده در تصویر جستجو می شود تا تطبیقی بین مدل و قسمتی از تصویر مشاهده شود. این امر با یک تابع شباهت^۵ برای تعیین میزان شباهت مدل با تصویر تخمین زده شده انجام می پذیرد. سپس، مدل موجود با استفاده از داده های موجود بروز رسانی می شود.

به طور کلی ردیابی بر مبنای مدل از سه مرحله اصلی ساخت مدل^۶، مطالعه ی نحوه ی تغییر اجزاء مدل و محدودیت های آنها^۷، و تکنیک های تخمین و روش های جستجو^۸ تشکیل شده است

۳-۲-۲- نحوه ی تطابق مشاهدات

از دید نحوه ی برقراری ارتباط بین اشیا در فریم های متوالی، روش های ردیابی را می توان به دو دسته ی روش های قطعی و روش های آماری تقسیم نمود.

الف) روش های قطعی

روش های قطعی برای به دست آوردن همبستگی بین نقطه های موجود در فریم N و $N+1$ از یک تابع هزینه که با استفاده از محدودیت های حرکت به دست می آید، استفاده می نمایند. هدف روش-

^۱ Model-based Humman Body Tracking

^۲ Analysis-by-Synthesis

^۳ Predict-Match-Update

^۴ Search Strategies

^۵ Similarity Function

^۶ Constraction of object model

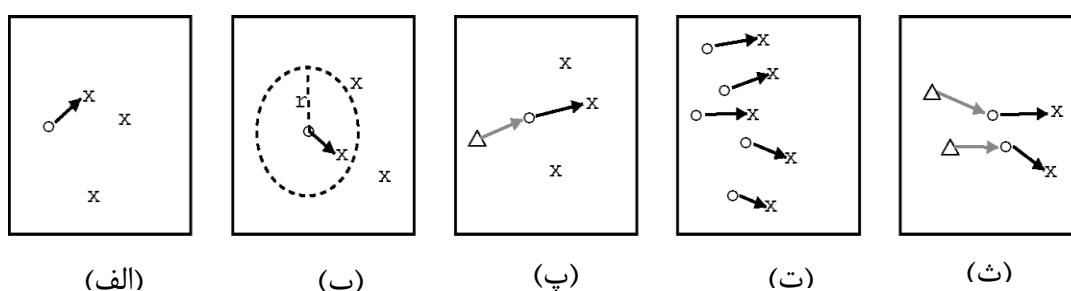
^۷ Representation of Prior Knowledge of Motion Models

^۸ Prediction and Search Strategies

های قطعی به حداقل رساندن تابع هزینه و مدل نمودن ردیابی به صورت یک مساله ی بهینه سازی می باشد.

هزینه ی تشابه عموماً توسط ترکیبی از محدودیت های زیر به دست می آید:

- مجاورت: فرض می نماید که مکان یک شی در دو فریم متوالی نمی تواند به طور قابل ملاحظه ای تغییر یابد. (شکل ۱-۳ الف)
- حداکثر سرعت: این محدودیت یک حد بالا برای سرعت شی در نظر گرفته و شباهت های موجود را به یک همسایگی دایره شکل محدود می نماید. (شکل ۱-۳ ب)
- تغییرات جزئی سرعت: این محدودیت فرض می نماید که سرعت و جهت یک شی نمی تواند به طور قابل ملاحظه ای تغییر یابد. (شکل ۱-۳ پ)
- حرکات مشترک: این شرط بر مبنای این اصل پایه گذاری شده است که تغییرات سرعت در یک همسایگی تا حد زیادی مشترک می باشد. این شرط برای تشخیص وابستگی بین اشیایی که با استفاده از چندین نقطه نمایش داده می شوند مورد استفاده قرار می گیرد. (شکل ۱-۳ ت)
- ثبات شکل^۱ (انعطاف ناپذیری): این اصل اشیاء را در یک فضای سه بعدی در نظر می گیرد، در نتیجه فاصله ی بین هر دو نقطه که مربوط به یک شکل باشند، باید ثابت بماند. این روش در ردیابی اهداف صلب کاربرد دارد. (شکل ۱-۳ ث)



شکل ۱-۳ عوامل محدودکننده در ردیابی قطعی افراد (الف) مجاورت (ب) حداکثر سرعت (پ) تغییرات جزئی سرعت (ت) حرکات مشترک (ث) ثبات شکل

^۱ Rigidity

ب) روشهای آماری

وجود نویز، مشاهدات ناقص و مشاهدات زاید موجب بروز عدم قطعیت در تشخیص و ردیابی افراد می‌گردند. به همین دلیل می‌توان ردیابی افراد را به صورت عبارات احتمالی مدل نمود. هدف از مدل احتمالی، مشخص کردن توزیع احتمال پسین مکان یک فرد یا حرکت وی بر اساس اندازه گیری‌های به دست آمده از تصویر می‌باشد.

اگر x_t نشان دهنده‌ی وضعیت^۱ یک فرد در زمان t (پارامتر نامعلوم در زمان فعلی) بوده و مشاهدات به دست آمده از دوربین در زمان t با z_t نشان داده شود، عمل ردیابی به صورت مساله‌ی محاسبه‌ی توزیع احتمال در دنباله‌ی $x_{1:t} = (x_1, \dots, x_t)$ مشروط به دریافت دنباله‌ی مشاهدات $z_{1:t} = (z_1, \dots, z_t)$ یعنی $p(x_{1:t}|z_{1:t})$ تعریف می‌گردد. بنابر قانون بیز خواهیم داشت :

$$p(x_{1:t}|z_{1:t}) = \frac{p(z_{1:t}|x_{1:t})p(x_{1:t})}{p(z_{1:t})} \quad (۱-۳)$$

در رابطه‌ی بالا $p(z_{1:t}|x_{1:t})$ شباهت^۲ نامیده می‌شود

۳-۳- فیلترینگ تصادفی^۳

فیلترینگ فرایند استخراج اطلاعات در مورد کمیتی مورد نظر در لحظه‌ی t با در دست داشتن اطلاعات اندازه گیری^۴ تا لحظه‌ی t و همچنین شامل لحظه‌ی t می‌باشد.

پیش بینی^۵ شکل پیشین^۱ فرایند تخمین است که هدف از آن پیش‌بینی توزیع کمیت مورد نظر در زمان $t + \tau$ و $(\tau > 0)$ توسط اطلاعات اندازه گیری تا لحظه‌ی t می‌باشد. پیش‌بینی، یک ناحیه جدید

^۱ State

^۲ likelihood

^۳ Stochastic Filtering Problem

^۴ Measured Data

^۵ Prediction

را در فضای تصویر یا فضای حالت ایجاد می کند و نیاز به جستجو کردن و پردازش در کل تصویر را حذف می کند. تخمین هر یک از پارامترهای مدل حالت، براساس مدل تغییرات این پارامترها در زمان است که می تواند سرعت، شتاب، ابعاد و یا سایر پارامترهای فرد باشد.

بروز رسانی^۲ شکل پسین^۳ برای فرایند تخمین است و هدف از آن بروز رسانی تخمین به دست آمده از مرحله پیش بینی است. بروز رسانی با در دست داشتن اطلاعات اندازه گیری در لحظه ی t انجام می شود.

مسئله ی فیلترینگ تصادفی در فضای حالت دینامیکی به شکل زیر در نظر گرفته می شود:

$$x(t) = f(t, x_t, v_t) \quad (۲-۳)$$

$$z(t) = h(t, x_t, n_t) \quad (۳-۳)$$

معادلات (۲-۳) و (۳-۳) به ترتیب معادله ی سیستم و معادله ی اندازه گیری نام دارند. x_t معرف بردار حالت و z_t معرف اندازه گیری و توابع f, h به صورت $f = r^{nx} \rightarrow r^{nz}$ و $h = r^{nx} \rightarrow r^{nz}$ تعریف می شوند که در حالت کلی دارای خاصیت تغییرپذیری با زمان هستند. n_t, v_t به ترتیب نویز پردازشی^۴ و نویز اندازه گیری^۵ می باشند.

به دلیل اینکه فیلتر کردن برای ردیابی افراد در حوزه ی زمان گسسته انجام می گیرد. معادلات (۳-۳)

(۲) و (۳-۳) را می توان به شکل زیر بازنویسی نمود:

$$x_k = f_k(x_{k-1}, v_{k-1}) \quad (۴-۳)$$

^۱ a prior

^۲ Updating

^۳ A posterior

^۴ Process Noise

^۵ Measurement Noise

$$z_k = h_x(x_k, n_k) \quad (۵-۳)$$

n_k و v_{k-1} نویز سفید با پارامترهای معلوم در حوزه‌ی زمان می باشند. معادله سیستم در (۳-۴)، معرف تابع توزیع احتمال گذر حالت^۱ به صورت $p(x_k|x_{k-1})$ و معادله‌ی اندازه‌گیری (۵-۳) معرف تابع شباهت به صورت $p(z_k|x_k)$ می باشد. هدف از فیلتر کردن، تخمین بهینه‌ی متغیر حالت در زمان با در دست داشتن اطلاعات اندازه‌گیری تا لحظه t ، شامل تابع توزیع اولیه $p(x_0)$ ، تابع توزیع احتمال گذر حالت $p(x_k|x_{k-1})$ و تابع شباهت $p(z_k|x_k)$ می باشد. مقدار تخمین‌زده شده به صورت $p(x_{0,k}|z_{0,k})$ یا $p(x_{0,k}|z_{0,k})$ خواهد بود که $x_{0,k}$ بردار حالت هدف از ابتدا تا لحظه‌ی جاری است.

۳-۳-۱- فیلترینگ بازگشتی بیزین غیر خطی^۲

قاعده بیز^۳ یک ابزار ریاضی برای استنتاج علت‌ها از یک سری مشاهدات می‌باشد. در ردیابی افراد موقعیت و یا سرعت هدف را می‌توان به صورت یک علت در نظر گرفته و با استفاده از پارامترهای مدل بیان نمود. سپس با دریافت مشاهدات مربوطه از سیستم می‌توان این علل را استنتاج نمود.

از دیدگاه بیزین عمل فیلترینگ تخمین بازگشتی متغیر حالت x_k در زمان k با در دست داشتن اطلاعات اندازه‌گیری تا زمان k یعنی $z_{1:k}$ می‌باشد. به عبارت دیگر کافی است در لحظه‌ی t توزیع $p(x_k|z_{1:k})$ را به دست آورد. در تئوری بیزین فرض بر این است که توزیع اولیه‌ی $p(x_0, z_0) = p(x_0)$ در دست است. تابع توزیع $p(x_k|z_{0:k})$ به طور بازگشتی طی دو مرحله‌ی پیش‌بینی^۴ و بروز رسانی^۵ قابل محاسبه است. جهت به دست آوردن فرمول‌های بازگشتی فرض می‌شود تابع توزیع احتمال مطلوب در زمان قبل یعنی در زمان $k-1$ به صورت $p(x_{k-1}|z_{1:k-1})$ محاسبه شده است. در این صورت

^۱ State Transition Probability Density Function

^۲ Recursive Nonlinear Bayesian Filtering

^۳ Bayes Rule

^۴ Prediction

^۵ Update

مرحله‌ی پیش‌بینی شامل استفاده از مدل سیستم $(\mathcal{Z}-3)$ جهت به دست آوردن تابع توزیع پیشین^۱ حالت در لحظه‌ی k با استفاده از معادله‌ی chapman-komogorov می‌باشد:

$$p(x_k | z_{1:k-1}) = \int p(x_k | x_{k-1}) \cdot p(x_{k-1} | z_{1:k-1}) dx_{k-1} \quad (6-3)$$

در بدست آوردن رابطه‌ی $(6-3)$ فرض شده است $p(x_k | x_{k-1}, z_{1:k-1}) = p(x_k | x_{k-1})$ می‌باشد که بیانگر داشتن خاصیت مارکف مرتبه اول برای توزیع گذر است. مدل احتمالی برای توزیع گذر حالت $p(x_k | x_{k-1})$ از معادله‌ی سیستم $(\mathcal{X}-3)$ و با در دست داشتن پارامترهای معلوم برای توزیع v_{k-1} قابل محاسبه است. بنابر تئوری بیزین خواهیم داشت:

$$p(x_k | z_{1:k}) = \frac{p(z_k | x_k) p(x_k | z_{1:k-1})}{p(z_k | z_{1:k-1})} \quad (7-3)$$

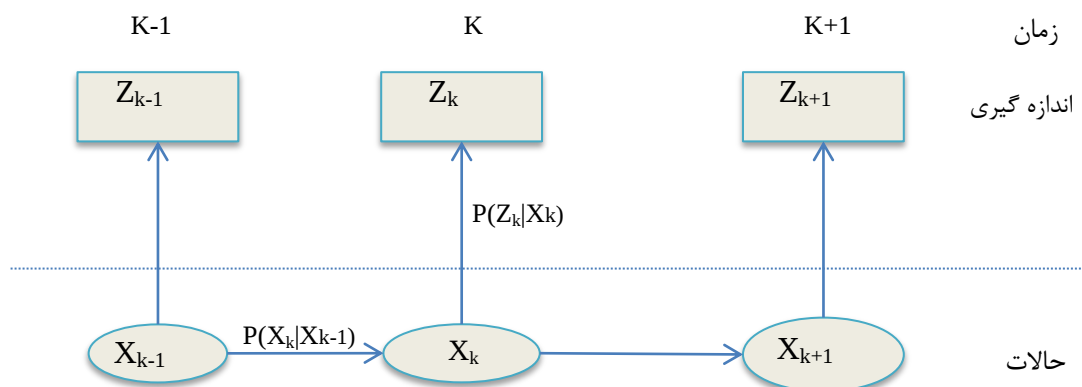
که:

$$p(z_k | z_{1:k-1}) = \int p(z_k | x_k) p(x_k | z_{1:k-1}) dx_k \quad (8-3)$$

عبارت $(8-3)$ وابسته به توزیع به دست آمده از مرحله‌ی پیش‌بینی و همچنین تابع شباهت است که توسط مدل اندازه‌گیری $(\mathcal{Z}-3)$ و با در دست داشتن پارامترهای توزیع n_k قابل محاسبه است. مراحل پیش‌بینی و بروز رسانی در شکل $(1-3)$ نشان داده شده است.

معادلات بازگشتی $(6-3)$ و $(7-3)$ پایه‌ی روش حل بهینه‌ی بیزین را تشکیل می‌دهند. در اغلب موارد بدست آوردن راه حل تحلیلی برای انتشار بازگشتی توزیع مطلوب امکان پذیر نیست. راه حل‌های تقریبی مختلفی جهت پیاده‌سازی فیلترینگ بیزین ارائه شده‌اند که بسته به ویژگی توزیع حالت هدف و نوع کاربرد مورد نظر قابل استفاده‌اند.

^۱ Aposteriori Estimate



شکل ۲-۳ مراحل پیش بینی و بروز رسانی در تئوری بیزین

برای یک شی متحرک در دنباله‌ای از تصاویر، بردار مکان شی را می‌توان به صورت بردار حالت $X = x_1, x_2, \dots, x_t$ نشان داد که در هر لحظه به یک مختصات (x, y) اشاره می‌کند. تغییرات این بردار در زمان با استفاده از رابطه زیر قابل محاسبه می‌باشد:

$$X_t = f^t(X_{t-1}) + W_t \quad (9-3)$$

که W_t نویز سفید می‌باشد. ارتباط بین مقدار اندازه‌گیری شده و بردار حالت توسط رابطه زیر محاسبه می‌گردد:

$$Z_t = h^t(X_t, N_t) \quad (3-10)$$

که N_t نیز نویز سفید جداگانه‌ای است.

برای تخمین بردار حالت X_t توسط تخمین تابع چگالی احتمال^۱ (PDF)، می‌توان طبق رابطه زیر عمل نمود:

$$p(X_t | Z_{1..t}) \quad (11-3)$$

معادله‌ی (۱۱-۳) را می‌توان به صورت بازگشتی با استفاده از فیلتر بیزین در دو مرحله‌ی تخمین^۱ و تصحیح^۲ حل نمود. در مرحله تخمین، فیلتر مقدار پیشین تابع PDF را با استفاده از بردار حالت جاری را طبق رابطه زیر به دست می‌آورد :

$$p(X_t | Z_{1,...,t-1}) \quad (12-3)$$

سپس در مرحله تصحیح، تابع شباهت^۳ $P(Z_t | X_t)$ به کار برده می‌شود تا مقدار پسین^۴ برای PDF طبق رابطه زیر به دست آید :

$$p(X_t | Z_{1,...,t}) \quad (13-3)$$

۳-۴- فیلتر کالمن

فیلتر کالمن یک ابزار ریاضی قدرتمند برای تخمین داده‌ها و ردیابی اشیا می‌باشد [۳۰]. این فیلتر حالت یک سیستم پویا را با استفاده از دنباله‌ی مشاهدات که می‌توانند ناقص باشند، تخمین می‌زند. فیلتر کالمن دارای این مزیت است که در آن بهره^۵ی فیلتر کالمن بسته به میزان نویز اندازه‌گیری و مدل حرکتی جسم ثابت نبوده و در بازه‌های زمانی مختلف به صورت بهینه محاسبه می‌شود. بنابراین این فیلتر از دقت بالایی برخوردار می‌باشد. تحقیقات بسیاری برای توسعه و بهبود فیلتر کالمن ارائه شده است.

برای تعیین موقعیت یک شیء، نیاز به اندازه‌گیری متغیرهای حالت آن مانند فاصله، سرعت و شتاب می‌باشد. در بیشتر مواقع امکان اندازه‌گیری برخی از این مقادیر ممکن نمی‌باشد. فیلتر کالمن قادر

Prediction - ۱

Correction - ۲

Like Hood - ۳

Posterior - ۴

^۵ Gain

است تنها با دریافت برخی از این مقادیر، سایر مقادیر را تخمین بزنند. می‌توان گفت در ردیابی افراد مهمترین کاربرد فیلتر کالمن، تخمین موقعیت شی، هنگام خروج شی از دید دوربین می‌باشد. در این موارد فیلتر کالمن بر اساس مشاهدات دریافت شده در زمان‌های قبل، قادر به تخمین مکان تقریبی شی می‌باشد. فیلتر کالمن در مواقعی که سیستم خطی بوده و نویز از نوع اضافه شونده گاوسی باشد بهینه است. در مقابل این فیلتر تنها حالاتی را در بر می‌گیرد که توزیع احتمال پارامترهای حالت تک قله‌ای^۱ باشد. فیلتر کالمن توانایی پیش بینی در مواردی از قبیل وجود همپوشانی بین اشیا و دینامیک پیچیده و غیر خطی اشیا (توزیع احتمال پارامترهای حالت، چند قله ای می باشد)، را ندارد. همچنین مسائلی که دارای ماهیت خطی و گاوسی نبوده و یا نویز آنها از نوع گاوسی اضافه شونده نباشد، با فیلتر کالمن نتایج مطلوبی را نخواهند داشت.

• تعریف ریاضی فیلتر کالمن

در فیلتر کالمن فرض می‌گردد که تابع توزیع پسین برای متغیر حالت در هر لحظه به شکل یک توزیع گاوسی بوده و توسط دو پارامتر میانگین و واریانس قابل توصیف می‌باشد. اگر توزیع $p(x_{k-1}|z_{1:k-1})$ به صورت گاوسی باشد توزیع $p(x_k|z_{1:k})$ نیز گاوسی خواهد بود به شرطی که:

- v_{k-1} و n_k از توزیع گاوسی با پارامترهای معلوم استخراج شده باشند.

- تابع $f(x_{k-1}, v_{k-1})$ و $h_k(x_k, n_k)$ معلوم بوده و تابعی خطی از متغیرهایش باشند.

بنابراین معادلات (۳-۳) و (۴-۳) را می‌توان به صورت ماتریس زیر نوشت:

$$x_k = f_k x_{k-1} + v_{k-1} \quad (۱۴-۳)$$

$$z_k = h_k x_k + n_k \quad (۱۵-۳)$$

h_k, f_k ماتریس معلوم و توابع خطی هستند. ماتریس کوواریانس بردارهای n_k, v_{k-1} ماتریس های P_k, Q_k می باشند. اگر n_k, v_{k-1} دارای میانگین صفر بوده و مستقل از هم باشند، می توان روابط بازگشتی را در الگوریتم فیلتر کالمن با استفاده از معادلات (۳-۵) و (۳-۶) به دست آورد.

برای تخمین حالت بازگشتی^۱ فیلتر کالمن یک سیستم اتفاقی^۲ را به صورت زیر در نظر گرفته می شود:

$$X_{k+1} = F(X_k, u_k, V_k) \quad (۳-۱۶)$$

که u_k ورودی سیستم و V_k ، نویز اضافه شده به محیط می باشد. هدف فیلتر کالمن تخمین حالت بردار X می باشد.

اگرچه X فقط از طریق مقدار اندازه گیری سنسور قابل دسترسی است اما این مقدار با نویز سنسور مخلوط شده است و مقدار واقعی X نمی باشد. این مقدار با رابطه زیر مشخص می شود:

$$Z_k = H(X_k, W_k) \quad (۳-۱۷)$$

W, V به ترتیب نویز سیستم و نویز سنسور می باشند که:

$$E[V_k] = 0, \quad E[V_k V_k^T] = Q_k, \quad \text{and} \quad E[V_i V_j^T] = 0, \quad \forall i \neq j \quad (۳-۱۸)$$

$$E[W_k] = 0, \quad E[W_k W_k^T] = R_k, \quad \text{and} \quad E[W_i W_j^T] = 0, \quad \forall i \neq j \quad (۳-۱۹)$$

تخمین حالت بازگشتی فرایندی تکرار پذیر^۳ است که طی آن بردار حالت x توسط اطلاعات موجود باز سازی می شود. این اطلاعات شامل مدل اندازه گیری و داده های دریافت شده می باشد.

^۱ Recursive State Estimation

^۲ stochastic

^۳ Iterative

$i \geq j$ و $(X_i|j)$ ، را تخمین حالت X_i می‌نامیم. با داشتن تخمین حالت $(X_k|k)$ و ورودی سیستم U_k می‌توان حالت $(X_{k+1}|k)$ را با استفاده از فرمول زیر تخمین زد:

$$X_{k+1|k} = E \left[f(X_k, u_k, V_k) | Z^k \right] \quad (۲۰-۳)$$

اگر برای رابطه (۳-۱۷) از نتایج روابط بالا استفاده شود، تخمین توزیع پیشین از مقدار اندازه گیری سیستم با رابطه ی (۳-۲۱) برابر خواهد شد.

$$Z_{k+1|k} = h(X_{k+1|k}, 0) \quad (۲۱-۳)$$

با مقایسه مقدار واقعی بردار Z_{k+1} و مقدار تخمین زده شده ی $(Z_{k+1|k})$ می‌توان به مقدار خطای اندازه گیری دست یافت. با اضافه کردن این خطا به مقدار تخمین حالت پیشین، می‌توان تخمین بهتری از مقدار واقعی داشت که به آن تخمین حالت پسین^۱ دست یافت:

$$X_{k+1|k+1} = X_{k+1|k} + K_{k+1} (Z_{k+1} - Z_{k+1|k}) \quad (۲۲-۳)$$

k_{k+1} بهره‌ی^۲ کالمن می‌باشد که برای مینیمم کردن کواریانس خطای حالت پسین استفاده می‌شود. برای بدست آوردن مقدار k از $(P_{k+1|k+1})$ نسبت به k مشتق گرفته و برابر صفر قرار می‌دهند (برای مینیمم کردن مقدار کوواریانس). مقداری که برای k بدست می‌آید مقدار بهینه ای برای رابطه ی (۳-۲۲) خواهد بود.

خطای تخمین حالت پیشین و پسین به صورت زیر محاسبه می‌شوند:

$$e_{k+1|k} = X_{k+1} - X_{k+1|k} \quad (۲۳-۳)$$

$$e_{k+1|k+1} = X_{k+1} - X_{k+1|k+1} \quad (۲۴-۳)$$

^۱ A posteriori

^۲ Gain

با توجه به مدل خطی در رابطه ی (۳-۱۷) تخمین حالت پیشین بدون نویز از رابطه زیر محاسبه

می شود

$$X_{k+1|k} = F_k X_{k|k} + u_k \quad (۳-۲۵)$$

خطای تخمین حالت پیشین، نیز از رابطه زیر محاسبه می شود:

$$e_{k+1|k} = F_k e_{k|k} + V_k \quad (۳-۲۶)$$

مقادیر z ، نیز از روابط زیر بدست می آید:

$$Z_{k+1|k} = H_{k+1} X_{k+1|k} \quad (۳-۲۷)$$

$$Z_{k+1} = H_{k+1} X_{k+1} + W_{k+1} \quad (۳-۲۸)$$

با استفاده از رابطه ی (۳-۲۴) می توان یک رابطه بازگشتی برای به دست آوردن خطای تخمین

حالت پسین بدست آورد:

$$e_{k+1|k+1} = e_{k+1|k} - k_{k+1} (H_{k+1} e_{k+1|k} + W_{k+1}) \quad (۳-۲۹)$$

مقدار کوواریانس خطا با توان دوم خطا بدست می آید:

$$P_{k+1|k} = E[e_{k+1|k} e_{k+1|k}^T] \quad (۳-۳۰)$$

$$P_{k+1|k+1} = E[e_{k+1|k+1} e_{k+1|k+1}^T] \quad (۳-۳۱)$$

با استفاده از روابط (۳-۲۶) و (۳-۲۸) می توان به رابطه ساده تری برای $(P_{k+1|k})$ دست یافت:

$$P_{k+1|k} = F_k P_{k|k} F_k^T + Q_k \quad (۳-۳۲)$$

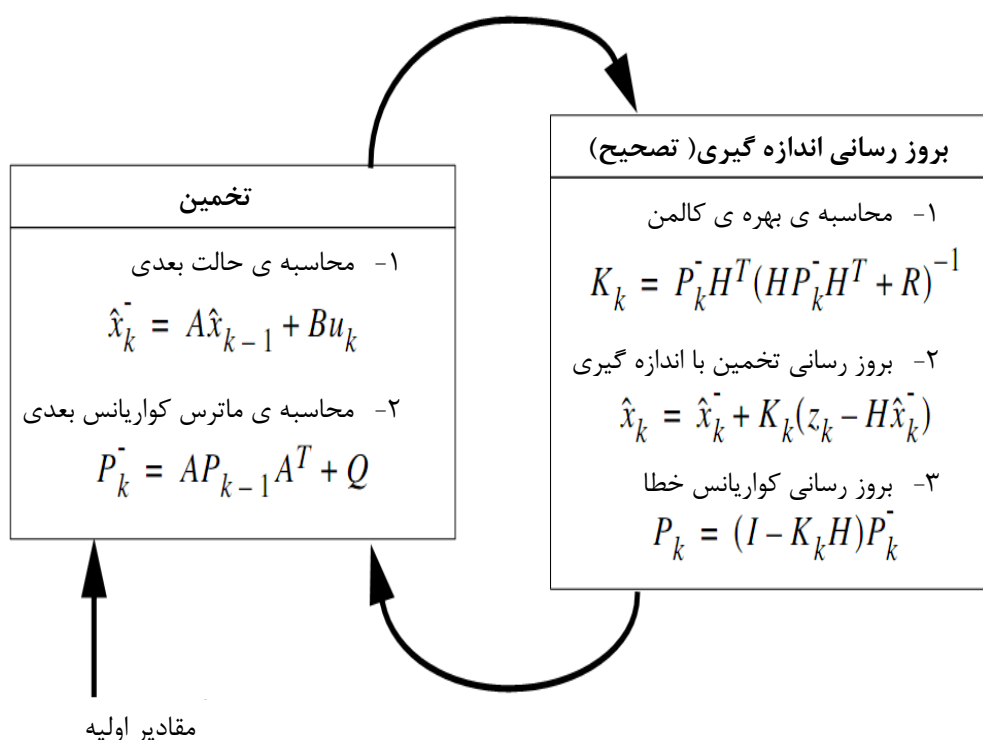
با جایگذاری رابطه (۳-۲۶) در رابطه (۳-۲۹) رابطه کوواریانس خطای پسین به صورت زیر خواهد

بود:

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k} H^T K^T - K H P_{k+1|k} + K (H P_{k+1|k} H^T + R) K^T \quad (۳-۳۳)$$

برای بدست آوردن مقدار k از $(P_{k+1|k+1})$ نسبت به k مشتق گرفته و برابر صفر قرار می دهند. (برای مینیمم کردن مقدار کوواریانس).

$$K = P_{k+1|k} H^T (H P_{k+1|k} H^T + R)^{-1} \quad (۳-۳۴)$$



شکل ۴-۲ مراحل فیلتر کالمن

روش ارائه شده برای فیلتر کالمن در صورت صادق بودن تمامی فرضیات در نظر گرفته شده، بهینه

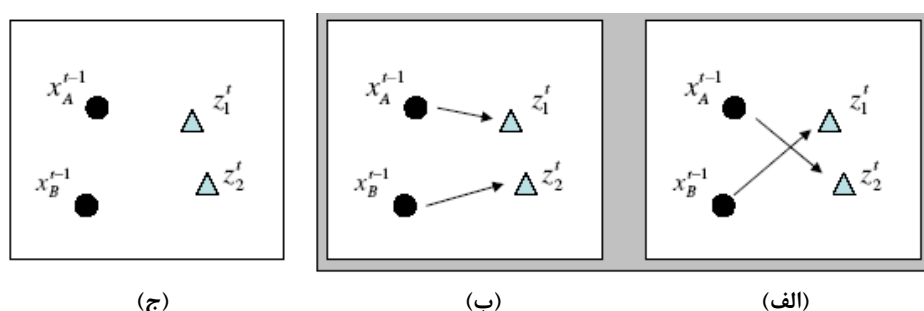
ترین حالت برای مسئله ی تخمین حالت می باشد. به این معنی که در صورت خطی و گاوسی بودن

توزیع حالت هیچ الگوریتمی منجر به جواب بهتر نخواهد داد. در برخی از کاربرد های عملی نمی توان

تمام فرض های لازم برای فیلتر کالمن را فراهم نمود. در این گونه مواقع می توان از الگوریتم های زیر اپتیمیم^۱ استفاده نمود. یکی از این الگوریتم ها الگوریتم کالمن توسعه^۲ [۳۱] است.

۳-۵- ردیابی اهداف چندگانه

مسئله ی ردیابی اهداف چندگانه^۳ (MTT) را علیرغم شباهت زیاد آن به ردیابی اهداف تکی، نمی-توان به عنوان تعمیمی ساده از مسئله ردیابی اهداف تکی به حساب آورد. دلیل این امر وجود مسئله ی موسوم به ارتباط اطلاعات است. برای وضوح بیشتر مسئله ی ردیابی دو هدف را در نظر بگیرید (شکل (۳-۳ الف)) فرض کنید در لحظه ی $t-1$ ، هدف A در مکان x_A^{t-1} آشکار سازی شده و به طور مشابه به وجود هدف B نیز در مکان x_B^{t-1} تثبیت شده است.



شکل ۳-۳ مثال ارتباط اطلاعات. الف) دو هدف (دایره ها) به همراه دو داده ی مشاهده شده (مثلث ها)، جهت ایجاد ارتباط. ب) حالت ممکن اول در مسئله ی ارتباط اطلاعات. ج) حالت ممکن دوم برای مسئله ی ارتباط اطلاعات.

در شکل (۳-۲ الف) اطلاعات Z_1^t و Z_2^t در زمان t توسط سیستم مشاهده می شوند. ابهام موجود در مسئله ی ردیابی اهداف چندگانه پاسخ به این سوال است که کدامیک از اطلاعات مشاهده شده توسط هدف A، و کدامیک از اطلاعات توسط هدف B تولید شده است. با فرض این که هر کدام از اهداف فقط یک داده ی اندازه گیری تولید کرده و هیچ آشکار سازی نادرست^۴ در سیستم صورت نگرفته

^۱ Sub optimal Algorithm

^۲ Extended Kalman Filter

^۳ Multiple-Target Tracking

^۴ False Alarm

است) عدم وجود پارازیت^۱ در محیط) دو حالت برای ارتباط اطلاعات مشاهده و مسیر حرکت^۲ می توان در نظر گرفت:

- Z_1^t متناظر با A و Z_2^t متناظر با B است (شکل ۳-۴ ب)

- Z_2^t متناظر با A و Z_1^t متناظر با B است. (شکل ۳-۴ ج)

اگر مسئله فوق را به حالت کلی N هدف با دقیقاً N داده‌ی مشاهده شده در هر لحظه و عدم وجود آشکارسازی نادرست تعمیم دهیم، تعداد حالت‌های ممکن برای تمام ارتباطات ممکن، برابر $N!$ خواهد بود که برای مقادیر بزرگ N، تعداد حالات برای کاربردهای عملی بسیار زیاد خواهد شد. به علاوه اگر تعداد حالت‌های ممکن را در طول پنجره‌ای به طول T در نظر بگیریم به عدد نمایی $(N!)^T$ دست خواهیم یافت. اگر فرضیات مربوط به تولید تنها یک داده‌ی اندازه گیری توسط هر هدف در هر واحد زمانی و همچنین عدم وجود آشکار سازی نادرست توسط سنسور در نظر نگیریم، در این صورت حجم محاسباتی بسیار بیشتر خواهد شد.

۳-۵-۱- فرمول بندی بیزین عمومی برای ردیابی اهداف چندگانه

ردیابی اهداف چندگانه با در نظر گرفتن مشکل ارتباط اطلاعات، حالت بعدی هر یک از اهداف را تخمین می‌زند. این مساله را می توان به صورت مسئله فیلترینگ بیزین متوالی^۳ با مشاهدات نویز آلود در نظر گرفت. با این تفاوت که در این حالت، با فضای حالت و فضای اطلاعات مشاهده‌ی پیچیده تری مواجه هستیم. در MTT متناظر با فضای حالت موجود در مسئله ردیابی اهداف تکی یعنی $X \in Z$ با فضای حالت متشکل از N هدف مواجه هستیم که قابل نمایش توسط مجموعه N تایی $(x_1, \dots, x_N) \in Z^N$ هستند. با توجه به اینکه عموماً تعداد اهداف برای سیستم نامعلوم است. فضای حالت توأم توسط رابطه زیر داده خواهد شد:

^۱ Clutter

^۲ Track

^۳ Sequential Bayesian Filtering Problem

$$S = Q \cup x \cup x^2 \cup x^3 \cup \dots \quad (3-35)$$

با توجه به رابطه فوق، فضای حالت کلی برابر اجتماع حالاتی است که هدفی موجود نباشد (ϕ) یک هدف داشته باشیم (x) و یا حالت های متشکل از تعداد محدودی هدف (x^N) باشد. مدل گذر از حالت چند هدفه ی s^{t-1} در لحظه ی $t-1$ به حالت s^t در لحظه ی t را می توان توسط احتمال گذر $p(s^t | s^{t-1})$ بیان کرد که مشابه معادله ی سیستم (دینامیک هدف) در ردیابی تک هدفه است؛ با این تفاوت که لازم است چگونگی ورود و خروج اهداف نیز مدل شوند. در ردیابی اهداف چند گانه اطلاعات اندازه گیری مجموعه ای از مشاهدات تولید شده توسط اهداف، و مشاهدات نادرست حاصل از وجود خطا در سنسورها می باشد. بنابراین اگر Z فضای اندازه گیری^۱ مربوط به یک هدف باشد فضای کلی اندازه گیری در حالت چند هدفه اجتماع تمام مجموعه های Z خواهد بود که آنرا با M نشان می دهیم. در اینصورت مدل اندازه گیری توسط یک تابع شباهت به صورت $P(m' | Z')$ ارائه خواهد شد که در آن m' مجموعه مشاهدات می باشد که مقادیر خود را از M خواهد گرفت. در مثال دو هدفه ی نشان داده شده در شکل (۳-۴) مجموعه مشاهدات به صورت $m' \in \{Z'_1, Z'_2\}$ می باشد. مسئله ارتباط اطلاعات به دلیل معلوم نبودن ترتیب اطلاعات موجود در مجموعه مشاهدات حاصل می شود.

از لحاظ تئوری می توان با اعمال مستقیم روش فیلترینگ بیزین بازگشتی استاندارد به مسئله عمومی ردیابی اهداف چندگانه مطرح شده در بخش قبلی، به توزیع فیلتر شده ی مطلوب $p(s | m^0, \dots, m')$ دست یافت. با این حال محاسبه توزیع فیلتر شده ی مطلوب در فضای حالت اهداف چندگانه s و در نظر گرفتن تعداد بسیار زیاد ترکیبات حاصل از فضاهای محتمل به واسطه مسئله ارتباط اطلاعات در عمل کاری بسیار مشکل و از لحاظ محاسباتی پیچیده است. بنابراین چالش اساسی موجود در تحقق سیستم های ردیابی اهداف چندگانه مدیریت پیچیدگی و حجم محاسباتی در سیستم با در نظر داشتن عملکرد مناسب برای عمل ردیابی است.

^۱ Measurement Space

روش های زیادی برای حل مشکل ارتباط اطلاعات در ردیابی اهداف چندگانه وجود دارد. یکی از این روش ها *MHT* می باشد که تمام حالات موجود برای مسیر موجود بین مشاهدات و اهداف را در نظر می گیرد. این باعث می شود که تعداد حالات ممکن با افزایش تعداد اهداف و تعداد مشاهدات به طور نمایی افزایش یابد. بنابراین این روش نیاز اساسی در کم کردن حجم اطلاعات دارد. از روش های دیگر حل این مشکل فیلتر استاندارد نزدیکترین همسایگی^۱ می باشد که هر کدام از اهداف را به نزدیکترین مشاهده در فضای هدف ارتباط می دهد. اما این روش باعث از بین رفتن یک سری از فرضیات امکان پذیر و شدنی می شود. فیلتر *JPDAF* یکی از روش های مناسبی است که برای حل مسئله ی ارتباط اطلاعات استفاده می شود.

۳-۵-۲- ردیابی فرضیات چندگانه^۲(MHT)

این روش تمام ارتباطات^۳ (موسوم به فرضیات) ممکن بین مشاهدات شامل مسیرهای موجود، مسیرهای جدید و مشاهدات نادرست را با در نظر داشتن محدودیت های ارتباطی به واسطه ناسازگاری های متقابل در نظر می گیرد. این روش نیازی به داشتن اطلاع از تعداد اهداف ندارد. در این روش مرحله اتخاذ تصمیم برای نحوه ی ارتباط اطلاعات به طور بهینه به منظور رسیدن اطلاعات مشاهده شده ی هر چه بیشتر تأخیر داده می شود. بنابراین لازم است فرضیات چندگانه ذخیره سازی شوند. الگوریتم *MHT* علیرغم داشتن توانایی بالا جهت ردیابی اهداف چندگانه با تعداد نامعلوم و مقاوم بودن در شرایط موجود مشاهدات نادرست با نرخ بالا دارای معایبی نیز هست. نیاز به حافظه بزرگ و افزایش نمائی بار حجم محاسباتی با افزایش تعداد اهداف از این جمله اند.

^۱ Nearest Neighbour Standard Filter

^۲ Multiple Hypothesis Tracking

^۳ Hypothesis

۳-۵-۳- الگوریتم JPDAF^۱

این روش شامل بروز رسانی هر یک از حالات مسیر توسط ترکیبات وزن دار تمامی مشاهدات می باشد که توسط فرتمن، شالم و شف^۲ مطرح شد. با توجه به اینکه بخش کلیدی در این روش محاسبه ی احتمال مرتبط بودن مشاهدات با مسیر هاست، این عمل با در نظر داشتن محدودیت های ارتباطی به واسطه ی ناسازگاری های متقابل در فرضیات صورت می پذیرد. از معایب این روش نیاز به معلوم بودن تعداد اهداف می باشد.

هر دو روش فوق تخمین هایی از توزیع فیلتر شده ی واقعی $p(s|m^0, ..., m')$ می باشند. روش MHT محاسبات بیشتری داشته و منجر به تخمین دقیق تری از توزیع می گردد. در مقابل JPDAF با به کارگیری اثر وزن دهی تمام مشاهدات، تصمیم گیری نرم تری برای ارتباط اطلاعات^۳ با هر مسیر داشته و همچنین محاسباتی کمتری نسبت به الگوریتم MHT نیاز دارد. در واقع الگوریتم JPDAF حالت خاصی از الگوریتم کلی تر MHT می باشد که روش خاصی جهت ترکیب فرضیات ایجاد شده توسط الگوریتم MHT و تولید یک فرضیه در هر زمان ارائه می نماید. در حالت کلی می توان گفت تمام روش های ارائه شده برای مسئله ی ارتباط اطلاعات، حالت خاصی از MHT هستند.

• ارتباط MHT و JPDAF

مثال ردیابی دو هدفه ی قبل را در نظر بگیرید. که در آن مسیر های A, B به طور مستقل و به ترتیب مطابق $p_A^{t-1}(x)$, $P_B^{t-1}(X)$ توزیع شده اند. داده های مشاهده در زمان t و z_1^t و z_2^t می باشند. برای سادگی فرض می شود مشاهدات نادرست در سیستم وجود ندارد. در این صورت براساس فرضیات الگوریتم MHT خواهیم داشت:

^۱ Joint Probability Data Association Filter

^۲ Fortman , Shalom , Scheffe

^۳ Soft Data Association

- H_0 : مسیر A با داده ی Z_1^t و مسیر B با داده ی Z_2^t مرتبط است.

- H_1 : مسیر A با داده ی Z_2^t و مسیر B با داده ی Z_1^t مرتبط است.

برای محاسبه احتمال ارتباط ها باید قبل از زمان t توزیع احتمال هر یک از مسیرها با توجه به

توزیع دینامیک اهداف بررسی گردد:

$$\hat{p}_j^{(t)} = \int_x p(x^t | x^{t-1}) p_j^{t-1}(x^{t-1}) dx^{t-1} \quad g \in \{A, B\} \quad (36-3)$$

می توان احتمال تولید مشاهدات z_1^t, z_2^t توسط مسیر های A, B را با استفاده از فرمول های زیر

محاسبه نمود:

$$\begin{aligned} \gamma_0 &= p(A \text{ generates } z_1^t \text{ and } B \text{ generates } z_2^t) \\ &= \int x^2 p(z_1^t | x_A) \cdot p(z_2^t | x_B) \cdot \hat{p}_A^t(x_A) \cdot \hat{p}_B^t(x_B) dx_A dx_B \end{aligned} \quad (37-3)$$

$$\begin{aligned} \gamma_1 &= P((A \text{ generates } z_2^t \text{ and } B \text{ Generates } z_1^t)) \\ &= \int x^2 p(z_2^t | x_A) \cdot p(z_1^t | x_B) \cdot \hat{p}_A^t(x_A) \cdot \hat{p}_B^t(x_B) dx_A dx_B \end{aligned} \quad (38-3)$$

با توجه به اینکه اطلاعات مشاهده ای مجموعه ی $\{Z_1^t, Z_2^t\}$ است. احتمال ارتباط ها به صورت زیر

محاسبه خواهد شد:

$$P(H_0) = \frac{\gamma_0}{\gamma_0 + \gamma_1}, \quad P(H_1) = \frac{\gamma_1}{\gamma_0 + \gamma_1} \quad (39-3)$$

احتمال ارتباطات تنها با استفاده از مدل اندازه گیری $P(Z|X)$ و توزیع های پیش بینی شده ی

مسیرها در رابطه (36-3) محاسبه شده است. در نتیجه دینامیک هدف، نقش تعیین کننده ای در

محاسبه ی احتمال ارتباطات خواهد داشت. با در نظر گرفتن هر یک از فرضیه ها حالت مسیرهای

A, B بر پایه قانون بیزین و با استفاده از مشاهده مربوط شده با آن به روز خواهد شد. در نتیجه، توزیع

مسیر A تحت فرضیه های H_0, H_1 به صورت زیر محاسبه خواهد شد:

$$P_A^t(X | H_0) = \alpha_0 \cdot P(Z_1^t | X) \cdot \hat{P}_A^t(X) \quad (40-3)$$

$$P_A^t(X | H_1) = \alpha_1 \cdot P(Z_2^t | X) \cdot \hat{P}_A^t(X) \quad (41-3)$$

در روابط فوق α_0, α_1 ثابت های نرمالیزسازی می باشند. حالت مسیر با مشاهدات مختلف تحت فرضیه های متفاوت به روز می شود؛ بنابراین الگوریتم *MHT* براساس هر یک از فرضیه ها حالت مسیر جداگانه ای را حفظ خواهد کرد. در صورتی که الگوریتم *JPDAF* فرضیه های چندگانه را با مخلوط کردن توزیع احتمال مسیرهای یکسان براساس تمام فرضیه ها ترکیب کرده و به فرضیه ی واحدی دست خواهد یافت که مفهوم تصمیم گیری نرم برای محاسبه ارتباط اطلاعات در هر مسیر می باشد.

$$P_{jpdaf,j}^t = P_j^t(x | H_0) \cdot P(H_0) + P_j^t(x | H_1) \cdot P(H_1) \quad , \quad j \in \{A, B\} \quad (42-3)$$

۳-۵-۴- سایر روش های ردیابی

علاوه بر فیلتر کالمن و مشتقات آن، روشهای دیگری نیز برای ردیابی اهداف ارائه شده اند. یکی از مشهورترین این روش ها فیلتر پارتیکل می باشد. ایده ی اصلی فیلتر پارتیکل اعمال بازگشتی فیلتر بیزی بر روی مجموعه ی نمونه هاست [۳۲]. این فیلتر به سیستم های خطی محدود نبوده و لزومی به محدود کردن نویز به نوع گاوسی اضافه شونده نیز وجود ندارد. روش فیلتر پارتیکل اخیراً در محیط های غیر گاوسی برتری هایی نسبت به روش فیلتر کالمن نشان داده است. اما پیچیدگی های محاسباتی بالای روش های مبتنی بر فیلتر پارتیکل مانع از کاربرد آنها در مسائل بزرگ نظیر ردیابی تعداد زیاد اشیاء متحرک می شود. به عنوان مثالی دیگر می توان به الگوریتم Condensation اشاره نمود [۳۳]. این الگوریتم توسعه ای از فیلتر پارتیکل می باشد که قادر به ردیابی توزیع های چند قله ای

نیز می‌باشد و با عناوینی نظیر فیلتر بوت استرپ بیزی^۱ و فیلتر مونت کارلو^۲ نیز شناخته می‌شود. این الگوریتم بر اساس نمونه برداری از تابع توزیع احتمال پسین تخمین زده شده در تصویر قبلی و پخش کردن این نمونه‌ها به منظور تشکیل تابع توزیع احتمال پسین برای تصویر فعلی عمل می‌کند. با این وجود از آنجائیکه این الگوریتم یک روش غیر پارامتریک است به تعداد نسبتاً زیادی نمونه جهت تضمین یک تخمین بیشترین احتمال^۳ صحیح از حالت فعلی نیاز دارد. الگوریتم دیگری در این زمینه با نام الگوریتم انتقال متوسط^۴ ارائه شده است [۳۴]. این الگوریتم خصوصاً در حضور شلوغی و بهم ریختگی زمینه دارای مزایایی نسبت به فیلتر کالمن می‌باشند.

^۱ Bayesian Bootstrap Filter^۲ Monte Carlo Filter^۳ Maximum Likelihood Estimation(MLE)^۴ Mean Shift

۳-۶- سیستم پیاده سازی شده

در این پایان نامه از فیلتر کالمن و ردیابی فرضیات چندگانه برای ردیابی افراد در تک تک دوربین-ها استفاده شده است. تنها اطلاعات لازم برای فیلتر کالمن در روش پیاده سازی شده، از اطلاعات زمانی و مکانی می باشند. موقعیت جسم، سرعت در راستای افقی و عمودی و اندازه ی شی و ویژگی-های هستند که می توانند در ماتریس حالت شی تاثیر بگذارند. برای برقراری ارتباط بین فرضیات موجود با مشاهدات دریافت شده از یک تابع هزینه به همراه الگوریتم ^۱SMM استفاده شده است. در صورتی که یک شی جدید با یکی از فرضیات موجود تطبیق داده شد، مشخصات آن فرضیه (ماتریس حالت و پارامترهای فیلتر کالمن) بروزرسانی می گردد. در صورتی که یک شی با فرضیات موجود منطبق نشود حالات متفاوتی قابل بررسی است. اولین حالت این است که واقعا یک شی جدید وارد سیستم شده است. دومین حالت می تواند ناشی از چند تکه شدن تصویر یک فرد به دلیل خطا در قطعه بندی تصویر باشد و در سومین حالت ممکن است یک شی که قبلا با شی دیگری هم پوشانی داشته دوباره ظاهر شده باشد. برای تشخیص حالت درست از یک تابع وزن دهی استفاده می شود که با استفاده از داده های موجود در پایگاه داده و دانش اولیه، به هر یک از حالات وزنی را اختصاص می دهد. به عنوان مثال هر چه شی در مجاورت مرزهای تصویر باشد، احتمال جدید بودن شی بیشتر می باشد.

در هر واحد زمانی، تمام اتفاقات رخ داده در قالب ساختمان داده ی استاندارد در پایگاه داده ذخیره می گردند. ذخیره ی رویدادها در پایگاه داده موجب کاهش پردازش های لازم برای تجزیه و تحلیل های بعدی و همچنین کاهش حجم داده ی ذخیره شده می گردد.

^۱ Stable Marriage Matching

هنگامی که فرد از نواحی خروجی سیستم عبور می‌نماید، سیستم ویژگی‌های مربوط به شی را استخراج نموده و در قالب رویداد خروج در پایگاه داده ذخیره می‌نماید. این اطلاعات در مرحله‌ی برقراری ارتباط بین مشاهدات در شبکه‌ی دوربین‌ها مورد استفاده قرار خواهند گرفت.

نتیجه‌گیری

در این بخش ابتدا تعریفی از ردیابی افراد در یک دوربین ارائه شده و چالش‌های آن مورد بررسی قرار گرفت. سپس برخی از روش‌های کلی برای ردیابی افراد معرفی شده و فیلتر کالمن به عنوان روش پیاده‌سازی شده در این پایان‌نامه به تفصیل شرح داده شد. همچنین الگوریتم MHT برای ردیابی اهداف چندگانه مورد بررسی قرار گرفته و در نهایت روش کلی پیاده‌سازی شده در این پایان‌نامه شرح داده شد.

الگوریتم ارائه شده در این پایان‌نامه تنها بر اساس اطلاعات حرکتی، افراد را ردیابی می‌نماید. همچنین به دلیل استفاده از فیلتر کالمن، امکان تشخیص هم‌پوشانی‌های معمول نیز فراهم آورده شده است. همچنین به دلیل عدم استفاده از ویژگی‌های ظاهری، افراد می‌توانند پوشش خود را تا حد قابل قبولی تغییر دهند.

فصل ۴: ردیابی اهداف در چند دوربین

اغلب سیستم های ردیابی برای پوشش نواحی نظارتی ، نیاز به شبکه ای از دوربین ها دارند. عمل ردیابی در شبکه ای از دوربین ها را می توان به صورت مسالهی برقراری ارتباط بین مشاهدات دریافت شده مربوط به یک فرد، در تک تک دوربین ها بیان نمود. در این فصل ردیابی افراد در چند دوربین مورد بررسی قرار خواهد گرفت. ابتدا روش های موجود برای ردیابی اهداف و مزایا و معایب هر یک مورد بررسی قرار خواهد گرفت. در ادامه راجع به ویژگی های قابل استفاده برای برقراری ارتباط بین مشاهدات توضیح داده خواهد شد. سپس یک ویژگی جدید بر مبنای ماتریس همجواری در فضای رنگی YCbCr برای برقراری ارتباط بین مشاهدات ارائه خواهد شد. این ویژگی علاوه بر استفاده از اطلاعات رنگی از اطلاعات بافتی ظاهر افراد نیز بهره برده و دارای ثبات و دقت بیشتری می باشد.

۴-۱- مقدمه

در اغلب موارد سیستم‌های ردیابی، برای نظارت بر محیط‌های بزرگ مانند فرودگاه‌ها، مراکز خرید و محیط‌های هوشمند نیاز است. در این سناریوها نمی‌توان با استفاده از یک دوربین تمام ناحیه‌ی نظارت را تحت پوشش قرار داد. این محدودیت می‌تواند ناشی از محدودیت‌های وضوح حسگر دریافت‌کننده‌ی تصاویر (دوربین) و یا موانع موجود در محیط باشد. بنابراین سیستم نظارتی یک ناحیه‌ی بزرگ، باید افراد را در شبکه‌ای از دوربین‌ها ردیابی نماید [۲]. همچنین هنگام استفاده از چند دوربین ممکن است پوشش دادن تمامی محدوده‌ی مورد نظر با استفاده از دوربین‌ها میسر نباشد. حتی در صورتی که بتوان تمام محیط را با استفاده از چندین دوربین با فضای دید مشترک تحت پوشش قرار داد، مشکلاتی نظیر کالیبره^۱ شدن دوربین‌ها با هم به وجود می‌آید و تغییر در فضای دید یک دوربین موجب تغییر در بقیه‌ی دوربین‌ها می‌گردد. بنابر این درسناپیوهای نظارتی^۲ تعریف شده در محیط-های واقعی نیاز به سیستم‌هایی است که بتوانند افراد را از طریق چندین دوربین با فضای دید غیر مشترک ردیابی نماید.

علاوه بر مسائلی که در ردیابی افراد در یک دوربین مطرح است، ردیابی با استفاده از چندین دوربین از جهات دیگر نیز چالش‌برانگیز است. مشاهدات دریافت شده از یک شی به صورت گسترده‌ای در زمان و مکان قابل توزیع می‌باشند و حضور افراد در زمان و مکان‌های مختلف در ناحیه‌ی دید دوربین‌ها تابع قانون خاصی نمی‌باشد. بنابر این نمی‌توان از ویژگی‌های زمان و مکان برای بررسی شباهت‌ها و مطابقت‌ها استفاده نمود. مشکل دیگر تغییرات به وجود آمده در ظاهر یک شی در محیط-های متفاوت است. ظاهر دریافت شده از یک شی، تابعی از نور محیط، ویژگی‌های هندسی^۳، نوع

^۱ Calibrate^۲ Surveillance^۳ Geometry properties

سطح^۱ شیء و مشخصات دوربین می‌باشد. بنابراین ظاهر یک شی در یک دوربین به دلیل عواملی از قبیل شرایط نوری محیط، موقعیت افراد، زاویه‌ی دید و سایر پارامترهای دوربین، می‌تواند با ظاهر همان شیء در دوربین دیگر متفاوت باشد.

ردیابی افراد در دو مرحله انجام می‌پذیرد. ابتدا افراد در تک تک دوربین‌ها ردیابی می‌شوند. پس از ردیابی افراد در هر یک از دوربین‌ها، برای ردیابی افراد در مجموعه‌ای از دوربین‌ها که هیچ فضای دید مشترکی با هم ندارند، نیاز به برقراری ارتباط بین مشاهدات مربوط به یک فرد می‌باشد. روش‌های متعدد انجام شده در این زمینه را می‌توان به سه گروه اصلی تقسیم بندی نمود.

اولین گروه که حجم عمده‌ای از تحقیقات را به خود اختصاص داده است، ویژگی‌های ظاهری را مبنای شناسایی افراد در دوربین‌های مختلف قرار می‌دهد. رنگ بارزترین ویژگی ظاهری هر فرد در سیستم‌های ردیابی افراد می‌باشد که معمولاً به صورت هیستوگرام رنگ مورد استفاده قرار می‌گیرد. با این وجود، استفاده از رنگ و مشتقات آن به دلیل حساسیت به شرایط نوری و پارامترهای دوربین باعث ایجاد محدودیت‌هایی در تشخیص و ارتباط صحیح بین مشاهدات می‌گردد [۳۵]. یکی از راه‌حل‌های ارائه شده برای حل این مساله، استفاده از تابع انتقال رنگ^۲ یا تابع انتقال روشنایی^۳ قبل از محاسبه‌ی میزان شباهت می‌باشد. به عنوان مثال [۳۶] از تابع انتقال روشنایی برای تبدیل روشنایی تصویر فرد در یک تصویر با تصویر دیگر استفاده نموده است. در [۳۷] به جای استفاده از هیستوگرام رنگ، رنگ‌های غالب تصویر مربوط به هر فرد استخراج شده و به عنوان معیاری برای برقراری ارتباط بین مشاهدات استفاده شده است. در [۳۸] توزیع رنگ به صورت تعدادی مدل گاوسی برای تعیین میزان شباهت استفاده شده و علاوه بر آن از یک تابع انتقال رنگ بر اساس نگاشت مدل‌های گاوسی برای بهبود کارایی بهره گرفته شده است. مقاله‌ی [۳۹] نیز ظاهر افراد را با استفاده از تعدادی توزیع

Object Surface^۱

Color Transfer Function^۲

Brightness Transfer Function^۳

گاوسی مدل نموده است. در این مقاله تعداد توزیع‌های گاوسی با استفاده از الگوریتم^۱ EM بهینه شده و برای کاهش اثر تغییرات نور، مقداری نویز به مدل رنگ مورد نظر اضافه شده است. مقاله‌ی [۴۰] یک تابع نگاشت تجمعی بر مبنای میانگین مقادیر رنگی ارائه داده و عنوان نموده است که توزیع تجمعی از نظر دقت نسبت به تابع توزیع استاندارد ارجحیت دارد.

دسته‌ی دوم شامل تحقیقاتی می‌گردد که از اطلاعات بیومتریک افراد مانند چهره، به عنوان معیاری برای تشخیص مشاهدات مربوط به یک فرد استفاده می‌نمایند. به عنوان مثال در [۴۱] از یک سیستم استریو برای استخراج عمق اشیا استفاده شده و در کنار آن از رنگ و چهره‌ی افراد برای برقراری ارتباط بین مشاهدات مربوط به یک فرد بهره گرفته 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 یک فرد توسط سیستم استریو. اگرچه نتایج ذکر شده توسط مولفین امیدوار کننده می‌باشند اما این سیستم از دو نظر قابل تامل می‌باشد. اول اینکه استفاده از ویژگی چهره، افراد را ملزم به حرکت در فاصله‌ی کمی از دوربین‌ها می‌نماید و دوم اینکه برای تشخیص صحیح چهره نیاز به تصاویری با کیفیت تقریباً بالا می‌باشد. در [۴۲] نیز ویژگی چهره به همراه رنگ و بافت لباس‌های افراد مورد بررسی قرار گرفته است. این سیستم بر روی یک پایگاه داده با ۱۱ فرد آزمایش شده است. نتایج این مقاله نشان می‌دهد ویژگی چهره با ۵۶٪ کمترین میزان دقت در بین سه ویژگی را دارد.

مقاله‌ی [۴۳] نحوه‌ی راه رفتن^۲ افراد را در هر یک از دوربین‌ها آنالیز نموده و به عنوان معیاری برای برقراری ارتباط بین مشاهدات مربوط به یک فرد در نظر می‌گیرد. این الگوریتم نیاز به پردازش‌های سنگین‌تری نسبت به سایر روش‌های ذکر شده دارد. همچنین نمونه‌های ویدئویی در شرایط کاملاً کنترل شده، در یک محیط و توسط یک دوربین تهیه شده‌اند.

^۱ Expectation Maximization

^۲ Gait

در نهایت الگوریتم‌های دسته‌ی سوم، از اطلاعات زمانی و مکانی استفاده نموده و مکان ورود افراد در دوربین بعدی را بر اساس معادلات احتمالی و حالات به دست آمده در دوربین فعلی تخمین می‌زنند. در مقاله‌ی [۴۴] مساله‌ی بازسازی مسیر افراد در چند دوربین، به صورت یک مساله‌ی بیزین تعریف شده است. سیستم پیاده‌سازی شده در این مقاله نیاز به مشخص نمودن دستی توپولوژی دوربین‌ها، مسیرهای ممکن و احتمال انتقالات بین دوربین‌ها می‌باشد. در [۴۵] نیز راهکاری برای تولید مسیر حرکت افراد در مجموعه‌ی دوربین‌ها با دیدهای مختلف ارائه شده است که برای مدل کردن دینامیک افراد از یک مدل مارکف استفاده می‌نماید. یکی از کارهای موفق در این زمینه در [۴۶] معرفی شده است. در این مقاله از ترکیب اطلاعات زمانی و مکانی به همراه ویژگی رنگ و تابع انتقال روشنایی برای برقراری ارتباط بین افراد استفاده شده است. این دسته از تحقیقات برای مواقعی مناسب می‌باشند که فاصله‌ی 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 دیگری انجام دهند سیستم دچار اختلال در تشخیص می‌گردد.

۴-۲- تعریف مساله

سیستمی با r دوربین داریم $C_1, C_2, C_3, \dots, C_r$ که فضای دید آنها با هم همپوشانی ندارد. همچنین فرض می‌نماییم n شی در محیط رفت و آمد می‌نمایند $P_1, P_2, P_3, \dots, P_n$ (تعداد محدود اما نامشخصی می‌باشد). هر یک از این اشیا در زمان‌های مختلف توسط دوربین‌های متفاوتی مشاهده می‌گردند. اگر مساله‌ی ردیابی افراد با استفاده از یک دوربین حل شده باشد و نتیجه‌ی آن $O_I = \{O_{I,1}, O_{I,2}, O_{I,3}, \dots, O_{I,m_I}\}$ مجموعه‌ی m_I مشاهده‌ی دریافت شده توسط دوربین I باشد، می‌توان مجموعه‌ی O را به عنوان مجموعه‌ی تمام مشاهدات دریافت شده توسط دوربین‌های مختلف در نظر گرفت. هر مشاهده‌ی $O_{I,a}$ توسط یک شی در زمان خاص و توسط یک دوربین دریافت شده است. هر مشاهده شامل چندین ویژگی از جمله ویژگی ظاهری شی $O_{I,a}(APP)$ و $O_{I,a}(ST)$ ویژگی

زمانی- مکانی (شامل مکان ، سرعت حرکت ، زمان و ...) می باشد به گونه ای که ویژگی های موجود در هر مشاهده مستقل از هم در نظر گرفته می شوند. مساله ی ردیابی با استفاده از چندین دوربین به- دست آوردن مجموعه ای از مشاهدات می باشد به گونه ای که این مشاهدات مربوط به یک فرد باشند [۲]. مساله ی گروه بندی مشاهدات را می توان به صورت مساله ی به دست آوردن زنجیره ی مشاهدات متوالی در نظر گرفت به گونه ای که مشاهده ی انجام پذیرفته توسط دوربین قبلی به مشاهده ی انجام پذیرفته توسط دوربین بعدی متصل گردد . خروجی سیستم، مسیر تردد هر یک از افراد در محیط می باشد. در این پروژه سه شرط محدود کننده ی اساسی، شامل فاصله ی ناچیز بین دوربین ها، وجود توالی بین دوربین ها و التزام افراد به حفظ شرایط حرکتی در نظر گرفته نشده است. به این معنی که فاصله ی بین دوربین ها می تواند قابل ملاحظه باشد. همچنین به دلیل عدم وجود توالی بین دوربین ها، ناحیه ی ورود بعدی نیز مشخص نمی باشد. همچنین فرد می تواند به صورت آزادانه و بدون در نظر گرفتن اطلاعات حرکتی در محیط حرکت نموده، یا حتی مکث نماید. به دلیل در نظر نگرفتن شرایط محدود کننده، عملاً پیش بینی زمان و مکان ورود فرد غیرممکن می باشد. بنابراین از $O_{I,a}(ST)$ که نشان دهنده ی ویژگی زمانی- مکانی شی (مکان ، سرعت حرکت ، زمان و ...) می باشند، صرف نظر می نماییم.

۴-۳- ویژگی های قابل استفاده برای برقراری ارتباط بین مشاهدات

همانگونه که در بخش نخست این فصل اشاره شد، ظاهر افراد به عنوان بارزترین ویژگی برای تشخیص افراد و برقراری ارتباط بین مشاهدات مربوط به یک فرد مورد استفاده قرار می‌گیرد. در این بخش برخی از ویژگی‌های قابل استخراج از ظاهر افراد مورد بررسی قرار خواهد گرفت.

۴-۳-۱- رنگ

رنگ، به عنوان یک ویژگی مهم در بینایی ماشین مطرح است که نمایش‌ها و تبدیلات فراوانی برای انجام کارهای متفاوت با آن ارائه شده است. این بخش نگاه اجمالی بر رنگ و فضاهای رنگی، به عنوان یک ویژگی برای برقراری ارتباط بین مشاهدات مختلف مربوط به افراد خواهد داشت. هدف از این بخش ارائه‌ی تمام فضاهای رنگی نبوده و تنها به معرفی فضاهای رنگی پرکاربرد در ردیابی افراد و همچنین فضای رنگ استفاده شده در این پایان نامه پرداخته شده است.

در صورت انتخاب رنگ برای توصیف ظاهر افراد، معمولاً از هیستوگرام رنگ و یا ترکیبی از توزیع رنگ‌های غالب استفاده می‌شود. هیستوگرام فقط به توصیف توزیع کلی رنگ شیء می‌پردازد و از جزئیات مکانی و چیدمان رنگ‌ها صرف نظر می‌کند.

ا) فضای رنگ RGB

RGB معمول‌ترین فضای رنگ برای استفاده در مانیتور و ذخیره‌سازی تصاویر است که قادر به نمایش ۱۶ میلیون رنگ می‌باشد. در این فضا، هر رنگ از ترکیب مقادیر سه رنگ اصلی قرمز، سبز و آبی به وجود می‌آیند. اگرچه فضای رنگ RGB برای نمایش تصاویر مناسب می‌باشد، اما از نظر مفهومی یکنواخت نمی‌باشد. به این معنی که تفاضل اقلیدسی در این فضای رنگ، با سیستم تفاوت رنگ‌ها در سیستم مفهومی انسان مرتبط نمی‌باشد. همچنین همبستگی زیادی بین ابعاد یک رنگ در فضای RGB وجود دارد. برخی محققان سعی نموده‌اند تا حد ممکن با استفاده از تبدیلاتی مشکلات

ذکر شده را رفع نمایند. یکی از ساده ترین این تبدیلات، تبدیل رنگ نرمال RGB می باشد که از فرمول زیر قابل محاسبه است.

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \quad (۱-۴)$$

ب) فضای رنگ HSV

فضای رنگ HSV یک بازنمایی از نقاط در فضای RGB است که سعی می نماید در ضمن پایین نگه داشتن پیچیدگی محاسبات، ارتباط مفهومی رنگ ها را به صورت دقیق تری مدل نماید. این مدل تا حد زیادی شبیه به سیستم درک رنگ توسط انسان ها می باشد. هر رنگ در این فضا توسط سه مولفه V^1, S^2, H^3 نمایش داده می شود. H توصیف کننده ی رنگدانه یا رنگ خالص بوده (مانند زرد خالص یا قرمز خالص) و نوع رنگ را بر حسب درجه بیان می کند. S میزان اشباع رنگ است و نشان می دهد که یک رنگ خالص تا چه حد توسط نور یا رنگ سفید رقیق شده است. V نیز نشان دهنده ی مقدار یا میزان روشنایی رنگ می باشد. برای تبدیل مقدار یک رنگ از فضای RGB به فضای HSV از فرمول های زیر استفاده می شود:

$$Max = \max(R, G, B) \quad (۲-۴)$$

$$Min = \min(R, G, B)$$

$$C = (max - min)$$

$$V = Max \quad (۳-۴)$$

$$S = \frac{C}{M}$$

Hue ¹
Saturation ²
Value ³

$$H' = \begin{cases} ((G - B) / C) \bmod 6 & , R = \max \\ 2.0 + (B - R) / C & , G = \max \\ 4.0 + (R - G) / C & , B = \max \end{cases} \quad (4-4)$$

$$H = H' * 60$$

ت) فضای رنگ YCbCr

این فضای رنگی برای نمایش تصاویر ویدیویی دیجیتال و همچنین کاربردهایی مانند فشرده‌سازی مورد استفاده قرار می‌گیرد. در این فضای رنگی Y معرف روشنایی پیکسل بوده و Cr و Cb به ترتیب نشان دهنده‌ی تفاضل رنگ قرمز و آبی از رنگ مرجع می‌باشند. برای تبدیل مقدار یک رنگ از فضای RGB به فضای YCbCr از فرمول زیر استفاده می‌نماییم

$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ 0.439 & -0.386 & 0.071 \\ -0.148 & -0.291 & 0.439 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5-4)$$

این فضای رنگی از چند جهت حائز اهمیت می‌باشد. اولاً در این فضا میزان روشنایی به صورت کامل از سایر مولفه‌ها جدا شده است. در حقیقت Y شامل فرکانس‌های لازم برای یک نمایشگر سیاه و سفید را فراهم می‌نماید. از طرفی چون مقدار روشنایی از مقادیر Cr و Cb کم شده است، این دو مولفه تا حد زیادی نسبت به تغییرات روشنایی پایدار هستند.

۴-۳-۲- ویژگی‌های مستقل از تغییر مقیاس^۱

ویژگی مستقل از تغییر مقیاس SIFT یکی از ویژگی‌های بسیار موفق در زمینه‌ی تشخیص اشیا می‌باشد که نسبت به تغییر مقیاس، انتقال، چرخش تصویر و تا حدودی نیز نسبت به تغییرات نور و تغییرات سه بعدی هندسی ثابت می‌باشند [۴۷]. در این روش ویژگی‌ها به صورت نقاط کلیدی تصویر مشخص می‌گردند. مزیت اصلی این روش نسبت به روش‌های ذکر شده پیشین در این است که

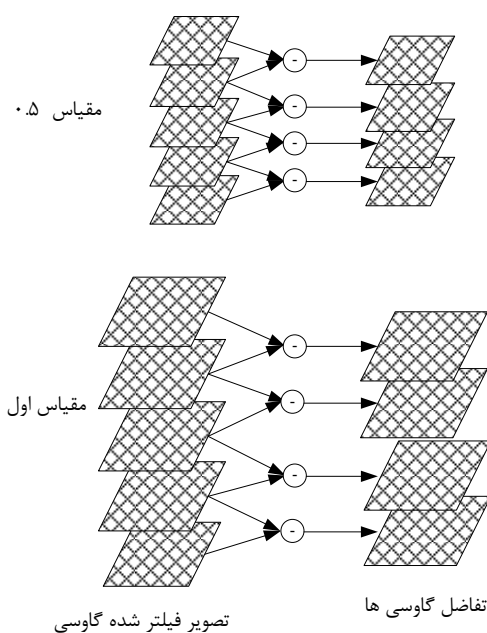
^۱ Scale Invariant Feature transform

ویژگی‌های محلی تصویر در برابر تغییرات مقیاس مقاوم بوده و حساسیت کمتری نسبت به نور و تغییرات هندسی سه بعدی دارند.

فرایند محاسبه‌ی ویژگی‌های پایدار با استفاده از تابع DOG^۱ و هرم تصویر یا هرم لاپلاسین انجام می‌پذیرد. هر سطح هرم از تفاضل نمونه مات شده^۲ از نمونه اصلی تصویر (DOG) حاصل می‌گردد. سطوح بعدی هرم نیز به روش بالا ولی از تصاویر کوچک شده ایجاد می‌شود. این فرایند در شکل ۴-۱ به نمایش درآمده است.

$$\begin{aligned} D(x, y, \delta) &= (G(x, y, k\delta) - G(x, y, \delta)) I(x, y) \\ &= L(x, y, k\delta) - L(x, y, \delta) \end{aligned} \quad (۴-۶)$$

$$L(x, y, \delta) = \frac{1}{2\pi\delta^2} \text{Exp}\left(-\frac{x^2 + y^2}{2\delta^2}\right) * I(x, y) \quad (۴-۷)$$



شکل ۴-۱: نحوه‌ی ایجاد هرم تفاضل گاوسی‌ها [۴۷]

^۱ Difference Of Gaussian

^۲ Blurred

با مقایسه هر نقطه‌ی نمونه در مقیاس فعلی با تمام همسایگی هایش در همان مقیاس (۸ همسایه سطح I، ۹ همسایه ی سطح i-1 و ۹ همسایه ی سطح i+1)، اکسترم‌ها در هرم DOG آشکار می‌گردند.

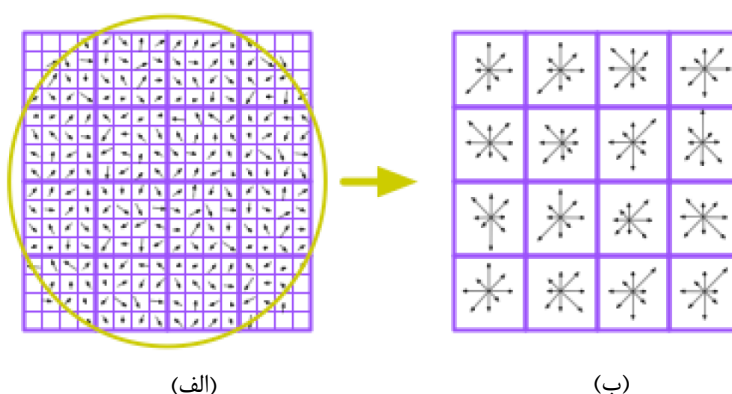
بعد از محاسبه‌ی نقاط کاندیدا، نقاطی که کنتراست پایینی داشته و نیز اکسترم‌های واقع بر روی لبه‌ها حذف می‌گردد. این نقاط موجب ناپایداری کلیدهای تصویری می‌گردد. در مرحله ی بعد با استفاده از فرمول زیر برای هر یک از کلیدهای به دست آمده زاویه ی خاصی اختصاص داده می‌شود.

$$f_x(x, y) = L(x+1, y) - L(x-1, y), f_y(x, y) = L(x, y+1) - L(x, y-1) \quad (۸-۴)$$

$$m = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad (۹-۴)$$

$$\theta = \tan^{-1} \left(\frac{f_y(x, y)}{f_x(x, y)} \right) \quad (۱۰-۴)$$

در پایان برای هر نقطه‌ی کلیدی، توصیف‌کننده‌هایی محاسبه می‌گردد. این توصیف‌کننده یک آرایه ی ۱۲۸ تایی از شدت گرادیان نقاط پیرامون کلید می‌باشد. شکل ۴-۲ نحوه ی ایجاد توصیف‌کننده را نمایش می‌دهد.



شکل ۴-۲ نحوه ی تولید توصیف‌کننده برای نقاط کلیدی در ویژگی مستقل از مقیاس
(الف) گرادیان در همسایگی نقطه ی کلیدی (ب) توصیف‌کننده

هر ویژگی SIFT بیانگر برداری از اندازه‌گیری‌های محلی تصویر می‌باشد به طوری که نسبت به انتقال تغییر مقیاس و چرخش ثابت بوده و در برابر تغییرات نور و تغییر شکل‌های محلی تصویر تقریباً ثابت می‌باشد. طبیعت محلی و چند مقیاسی بودن این ویژگی آن را نسبت به نویز، شلوغی و انسداد غیر حساس می‌سازد.

محققان در سالهای اخیر به دلیل ثبات این ویژگی نسبت به تغییرات روشنایی، چرخش و مقیاس اشیاء، از آن در بسیاری از کاربردهای شناسائی اشیاء استفاده کرده‌اند که منجر به نتایج بسیار خوبی گردیده است. به عنوان مثال کارهایی در زمینه ی تشخیص چهره، با استفاده از این ویژگی انجام شده است [۴۸]. برای تضمین نتیجه‌ی مناسب توسط این الگوریتم شرایطی نیز وجود دارد. اولین شرط ثابت ماندن شکل کلی نقاط متناظر در دو تصویر است، شبیه آنچه در نقاط مربوط به یک جسم صلب برقرار می‌باشد. نقاط متناظر موجود در بدن فرد بنا به ماهیت روان بودن و تغییر پذیری لباس افراد از این ویژگی برخوردار نمی‌باشد. همچنین وضوح و کیفیت تصویر دریافتی در سیستم‌های نظارتی در اغلب موارد پایین می‌باشد. استفاده از نمونه‌هایی با وضوح بالا با توجه به حجم بسیار بالای محاسبات انجام یافته در مراحل مختلف یک سیستم ردیابی و ماهیت بیدرنگ و یا نزدیک به بیدرنگ سیستم-های نظارتی عملاً غیرممکن می‌باشد. بیشتر سیستم‌های نظارتی موجود برای پیاده‌سازی الگوریتم خود از تصاویری با وضوح 320×240 و یا حداکثر 640×480 استفاده می‌نمایند. در نتایج حاصل از نمونه‌های استفاده شده در این پایان‌نامه، متوسط ابعاد یک فرد در نمونه‌ای که افراد نزدیک به دوربین حرکت می‌نمایند، از 75×180 تجاوز نمی‌نماید. حال آنکه برای تشکیل هرم تغییرات گاوسی به منظور تشخیص نقاط کلیدی، این الگوریتم تصاویر را تا سه مرتبه (در الگوریتم پیشنهاد شده در الگوریتم اصلی) تغییر سایز داده و کوچک‌تر می‌نماید. این عمل باعث از دست رفتن اغلب اطلاعات مفید و کم شدن تعداد نقاط کلیدی می‌گردد. در پیاده‌سازی‌هایی انجام پذیرفته در این پروژه، متوسط تعداد نقاط کلیدی برای هر فرد حدود ۷ نقطه‌ی کلیدی در تصویر مربوط به هر فرد به دست آمد. نسبت نقاط

کلیدی مربوط به یک فرد، که امکان انطباق داشتند، حتی در یک دوربین، بسیار کم و نزدیک به دو نقطه‌ی کلیدی بود. بنا به دلایل ذکر شده، استفاده از این ویژگی برای برقراری تناظر بین مشاهدات مربوط به یک فرد مناسب نمی باشد. نتایج به دست آمده از آزمایشات نیز این امر را تایید می نمایند.

۴-۴- معیار های اندازه گیری شباهت

پس از استخراج ویژگی از مشاهدات به دست آمده از تصاویر، نیاز به معیاری برای مقایسه‌ی ویژگی‌های به دست آمده می باشد. معیار شباهت نقش بسزایی در کارائی سیستم دارد و بسته به نوع داده ها باید از معیار شباهت مناسب استفاده کرد. ساده ترین معیار فاصله، فاصله‌ی اقلیدسی می باشد. برای دو بردار q و p که از k جزء تشکیل شده است از رابطه‌ی زیر استفاده می شود.

$$e(Q, P) = \sum_{i=1}^k \sqrt{(q_i - p_i)^2} \quad (۴-۱۱)$$

معیار محاسبه‌ی دیگر ضریب واگرایی **Kullback-Leibler** می باشد. مقدار شباهت دو بردار از طریق رابطه‌ی زیر محاسبه می گردد. P عموماً نشان دهنده‌ی بردار ویژگی برای مقدار اصلی و Q نشان دهنده‌ی مقدار تخمین زده شده و یا تشخیص داده شده می باشد. یکی از معایب این روش غیر متقارن بودن آن می باشد.

$$k(Q, P) = \sum_{i=1}^k p_i \log \frac{p_i}{q_i} \quad (۴-۱۲)$$

تابع Jensen معیار مناسب دیگری است که می توان برای مقایسه‌ی دو بردار استفاده نمود. مقدار این تابع بین صفر و یک می باشد. برای دو بردار کاملاً یکسان، خروجی این تابع برابر با صفر و برای دو بردار کاملاً متفاوت عدد یک می باشد. این تابع از طریق روابط زیر به دست می آید.

$$j(p, q) = \frac{1}{2} \sum_{i=1}^k \{P'_i \log_2 P_i + Q'_i \log_2 Q_i - (P'_i + Q'_i) \log_2 (P_i + P_i)\} \quad (۱۳-۴)$$

$$P' = \frac{P_i}{\sum_{i=1}^k p_i}, Q' = \frac{Q_i}{\sum_{i=1}^k q_i} \quad (۱۴-۴)$$

ضریب Bhattacharyya معیار دیگری است که برای مقایسه ی دو بردار نرمال (ماکزیمم مقدار یک) می توان از آن استفاده نمود. در حقیقت این ضریب مشخص کننده ی میزان هم پوشانی دو نمونه ی آماری می باشد. اگر هیچ اشتراکی بین دو بردار نباشد، مقدار این ضریب صفر می گردد. این ضریب با استفاده از رابطه ی زیر به دست می آید.

$$BC(P, Q) = \sum_{i=1}^n \sqrt{\left(\sum p_i \cdot \sum q_i \right)} \quad (۱۵-۴)$$

معیار دیگر که با ضریب Bhattacharyya در ارتباط است، فاصله ی Hellinger که به صورت عکس ضریب Bhattacharyya در نظر گرفته می شود.

$$H(P, Q) = \sqrt{1 - BC(P, Q)} \quad (۱۶-۴)$$

۴-۵- ویژگی معرفی شده در پایان نامه

در بیشتر تحقیقات انجام پذیرفته، ویژگی‌های استخراج شده از ظاهر افراد برای تمیز بین اشیاء موجود در یک سیستم نظارتی استفاده شده است. رنگ و بافت دو ویژگی سطح پایین هستند که عموماً برای طبقه‌بندی اشیاء مورد استفاده قرار می‌گیرند. همانگونه که در قسمت‌های قبل ذکر شد، رنگ، یک ویژگی بارز در توصیف ظاهر افراد می‌باشد. روش معمول برای مقایسه‌ی افراد با استفاده از رنگ، استخراج هیستوگرام رنگ و تطبیق آن در مشاهدات مختلف می‌باشد. هیستوگرام رنگ، تنها توزیع فراوانی مقدار رنگ را اندازه‌گیری نموده و از اطلاعات مکانی مقادیر رنگ صرف‌نظر می‌نماید. همچنین به دلیل حساسیت رنگ نسبت به نور سیستم ردیابی با استفاده از هیستوگرام رنگ، کارایی پایینی خواهد داشت. در مقابل بافت تصویر یک شیء با صرف‌نظر کردن از اطلاعات رنگی، سعی در توصیف ویژگی‌های مکانی شیء می‌نماید. ماتریس هم‌جواری سطح خاکستری^۱ یک روش شناخته شده برای استخراج بافت شکل می‌باشد [۴۹].

در این پایان نامه، یک ویژگی جدید برای برقراری ارتباط بین مشاهدات مربوط به افراد پیشنهاد شده است. اساس کار الگوریتم پیشنهادی، استخراج بردار ویژگی از ماتریس هم‌رخداد^۲ رنگ در فضای رنگ YCbCr می‌باشد. بردار ویژگی معرفی شده علاوه بر استفاده از اطلاعات رنگ، تا حدی از اطلاعات بافت تصویر نیز بهره می‌گیرد. این ویژگی به چند دلیل نسبت به هیستوگرام رنگ دارای ثبات بیشتری می‌باشد. اولاً استفاده از ماتریس هم‌رخداد با شمارش مقادیر رنگ پیکسل‌های همجوار، علاوه بر فراوانی مقادیر رنگ، تا حدی اطلاعات مکانی آنها را نیز در نظر می‌گیرد. برای غلبه بر مشکل حساسیت رنگ به نور ابتدا تصویر از فضای رنگ RGB به فضای رنگ YCbCr تبدیل شده است. همانگونه که در بخش فضاهای رنگی اشاره شد، فضای رنگی YCbCr به دلیل جداسازی مولفه-

^۱ Gray Level Co-Occurrence Matrix

^۲ Co-occurrence Matrix

ی روشنایی، نسبت به تغییرات نور، دارای ثبات بیشتری از توصیف رنگی RGB می باشد. در ادامه ی این فصل ابتدا ماتریس همجواری معرفی شده و سپس نحوه ی استخراج بردار ویژگی توضیح داده می شود.

۴-۵-۱- ماتریس همرخداد

ماتریس هم رخداد، مشخص کننده ی فراوانی رخداد مقادیر خاکستری دو پیکسل وابسته از تصویر است که در فاصله و جهت بخصوصی از یکدیگر قرار گرفته اند. ماتریس همرخداد G_{ij} ماتریس مربعی است که مقدار عنصر I_i آن، نشان دهنده فرکانس نسبی رخداد دو پیکسل در فاصله و زاویه ی خاص، یکی با شدت خاکستری i و دیگری با شدت خاکستری j می باشند [۵۰]. زاویه های ممکن بین دو پیکسل ۰، ۴۵، ۹۰ و ۱۳۵ درجه می باشد. تعداد سطوح شدت ممکن در تصویر، مشخص کننده ی ابعاد ماتریس G می باشد. برای تصویر ۸ بیتی، اندازه G برابر 256×256 خواهد بود. به منظور کاهش بار محاسباتی و همچنین تا حدی از بین بردن اثر نور، از کمی کردن شدت^۱ رنگ استفاده می شود. از ماتریس هم رخداد عموماً برای مشخص نمودن اطلاعات بافتی استفاده می شود. در [۵۱]، ۱۴ ویژگی از جمله همبستگی، کنتراست، یکنواختی و آنتروپی، بر اساس مقادیر این ماتریس تعریف شده است. مزیت استفاده از ماتریس همرخداد برای نمایش بافت تصویر، این است که تا اندازه ای موقعیت مکانی پیکسل ها نیز در نظر گرفته می شود.

۴-۶- استخراج ویژگی

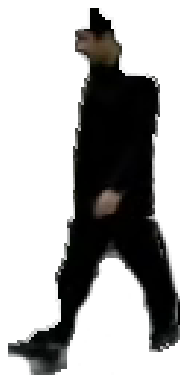
ماتریس همرخداد سطح خاکستری، اطلاعات رنگی شی را نادیده می گیرد. ویژگی تعریف شده در این پایان نامه یک ماتریس همرخداد به صورت جداگانه برای هر یک مولفه های رنگی Y , Cb , Cr محاسبه می گردد. همانگونه که در شکل ۴-۳ نشان داده شده است، ماتریس هم رخداد یک ماتریس

^۱ Quantization

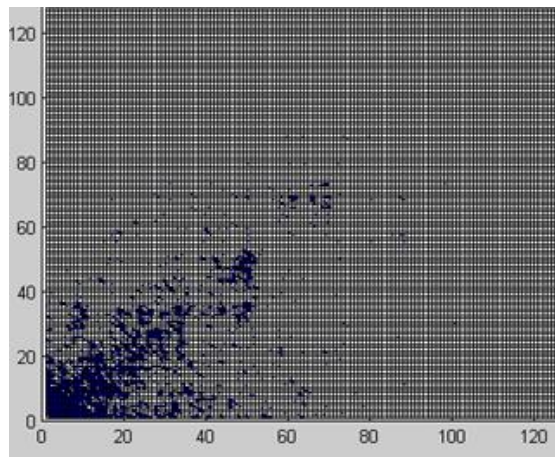
مربعی است که برای افراد بیشتر مقادیر آن صفر و یا بسیار کوچک هستند. با توجه در اشکال می توان دریافت که مقادیر این ماتریس با دور شدن از قطر اصلی به صفر میل می نماید. هر چه تصویر دارای تعداد رنگهای بیشتری باشد که به صورت تصادفی توزیع شده باشند، احتمال یکنواخت بودن ماتریس هم‌جواری بیشتر خواهد بود. در مقابل هرچه تعداد رنگ های یک تصویر کمتر بوده و پیکسل های یک‌رنگ در یک ناحیه باشند، مقادیر ماتریس هم‌جواری متمایل به قطر اصلی خواهد بود.

از آنجایی که پوشش افراد معمولاً از تعداد محدودی رنگ تشکیل شده است که به صورت یکنواخت پراکنده شده است، بنابراین مقادیر غیر صفر ماتریس هم‌رخداد، برای تصاویر استخراج شده ی افراد حول قطر اصلی پراکنده شده‌اند. می‌توان از سایر مقادیر ماتریس صرف‌نظر نموده و عناصر روی قطر اصلی را به عنوان ویژگی مناسب برای مقایسه ی مشاهدات در نظر گرفت.

پس از محاسبه‌ی بردار مربوط به مقادیر قطر اصلی ماتریس هم‌جواری، برای از بین بردن تاثیرات اندازه‌ی تصویر، آن را به وسیله‌ی تقسیم مقادیر آن بر مقدار بیشینه‌ی بردار، نرمال می‌نماییم. نتیجه برداری است که همواره مقدار ماکزیمم آن یک است. پس از نرمال سازی برای بهبود کارائی، توزیع به دست آمده را تبدیل به توزیع تجمعی می‌نماییم. شکل ۴-۴ نشان دهنده ی بردار های به دست آمده برای نمونه های موجود می باشد.



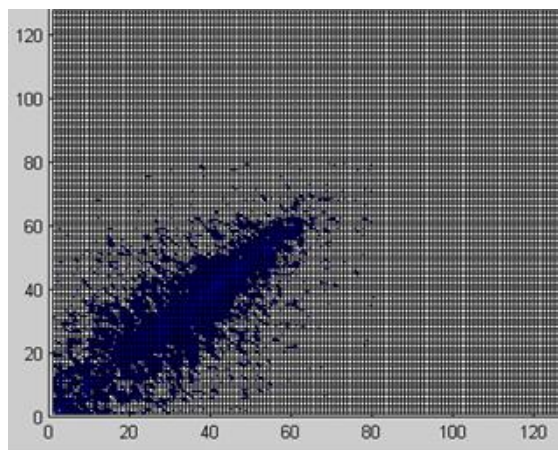
(الف)



(ب)



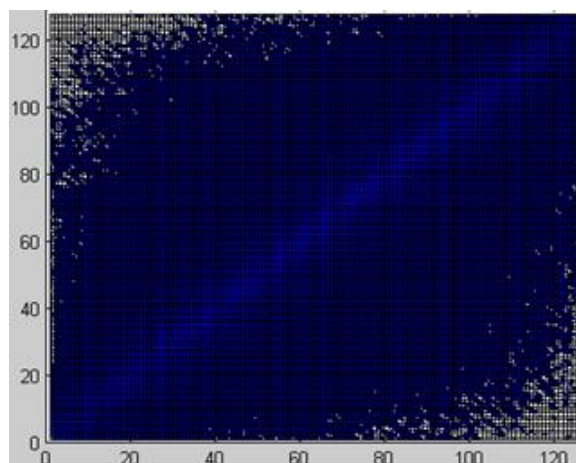
(پ)



(ت)



(ث)



(ج)

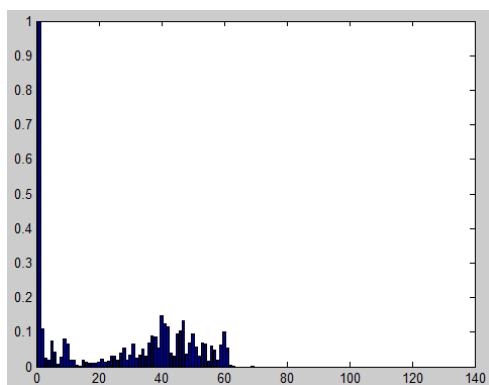
شکل ۳-۴ تصاویری با بافت های مختلف و ماتریس هم رخداد متناظر آنها



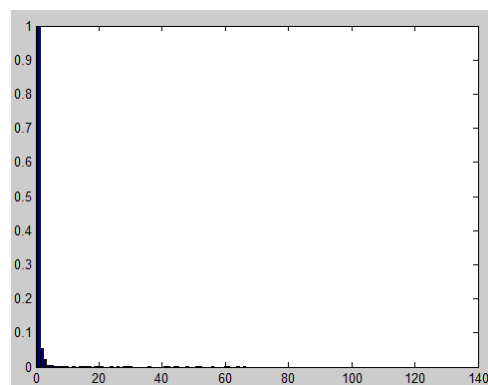
(الف)



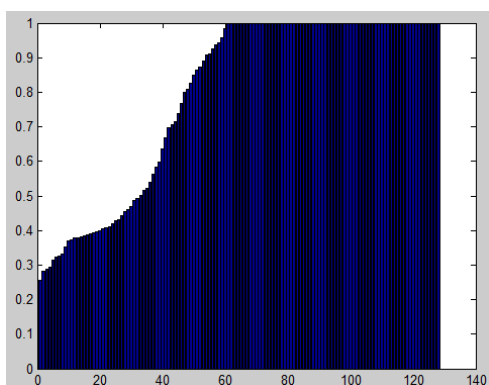
(ب)



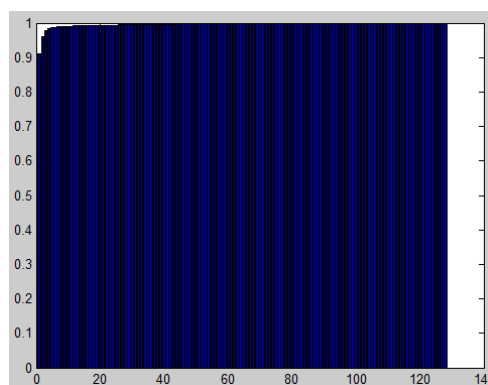
(پ)



(ت)



(ث)

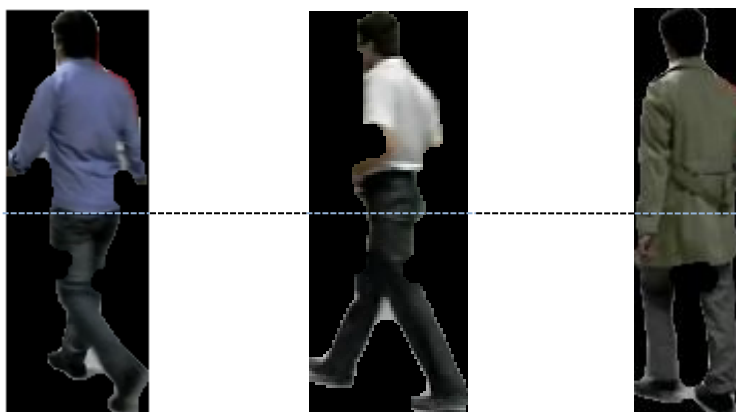


(ج)

شکل ۴-۴ تصویر دو فرد و بردارهای به دست آمده با استفاده از روش پیشنهادی

۴-۷- بهبود کارایی با تقسیم ظاهر افراد به چند قسمت

پوشش افراد را می‌توان به قسمت‌های سر، بالا تنه و پایین تنه تقسیم نمود. افراد می‌توانند برای هر یک از این قسمت‌ها از پوشش‌های متنوعی استفاده نمایند. در برخی از موارد ترکیب‌های مختلف می‌توانند منجر به تولید هیستوگرام یکسانی گردند. به عنوان مثال فردی با پیراهن سفید و شلوار مشکی، تقریباً دارای هیستوگرامی شبیه به فردی با شلوار سفید و پیراهن مشکی می‌باشد. این قضیه در مورد ویژگی معرفی شده در این پایان نامه نیز تا حد زیادی صادق می‌باشد. از این رو می‌توان کارایی سیستم را با استفاده از تقسیم ظاهر فرد به چند بخش بهبود داد. در سیستم ردیابی پیشنهاد شده ظاهر افراد به دو قسمت بالاتنه و پایین تنه تقسیم شده و ویژگی‌های به دست آمده به طور محلی با هم مقایسه شده و نتیجه از ترکیب نتایج محلی به دست می‌آید. نیمه‌ی بالایی شکل به عنوان بالاتنه و نیمه‌ی پایینی به عنوان پایین تنه در نظر گرفته شده است. به دلیل اندازه‌ی کوچک تصویر مربوط به هر فرد و همچنین وضوح پایین دوربین‌ها از تقسیم قسمت سر صرف‌نظر شده و این بخش به عنوان قسمتی از بخش بالاتنه در نظر گرفته شده است. شکل ۴-۵ محدوده‌ی هر قسمت را نمایش می‌دهد.



شکل ۴-۵ محدوده‌ی تقسیم ظاهر افراد

۴-۸- تطبیق ویژگی ها

ویژگی های تعریف شده در منطقه ی ورود و خروج محاسبه شده و در پایگاه داده ذخیره می شوند. هنگامی که فردی در ناحیه ی ورود یک دوربین قرار می گیرد، ویژگی های آن استخراج شده و با ویژگی های موجود در پایگاه داده مقایسه می گردد. در صورتی که سیستم بتواند ویژگی جدید را با یکی از مقادیر موجود تطبیق دهد، اطلاعات فرد مربوطه شامل برچسب و شماره ی دوربین قبلی در پایگاه داده بروز رسانی می گردد. معیار های ویژگی متنوعی وجود دارند که برخی از آنها در بخش ۴-۴ معرفی شده اند. نتایج نشان می دهد فاصله ی جنسن کارآمد ترین معیار برای ویژگی معرفی شده در این پایان نامه می باشد.

۴-۹- نتیجه گیری

در این بخش ابتدا تعریف کلی از مساله ارائه شده و برخی از راه حل های موجود برای حل مساله ی ردیابی افراد در شبکه ای از دوربین ها با دیدهای مجزا مطرح گردید. سپس رنگ به عنوان یکی از مهمترین ویژگی های موجود برای ردیابی افراد مطرح شده و برخی از فضاها ی رنگی مورد بررسی قرار گرفت. به منظور تعیین معیاری برای مشابهت دو ویژگی، برخی از روش های موجود مطرح شد.

سپس، ویژگی جدیدی معرفی گردید که دارای ثبات بیشتری نسبت به ویژگی های استخراج شده از هیستوگرام رنگ دارد. این ویژگی عناصر قطره های اصلی ماتریس هم جواری را استخراج نموده و پس از نرمال سازی به یک توزیع تجمعی تبدیل می نماید. در نهایت به منظور بالا بردن کارایی، تغییرات جزئی اعمال شده در الگوریتم توضیح داده شد.

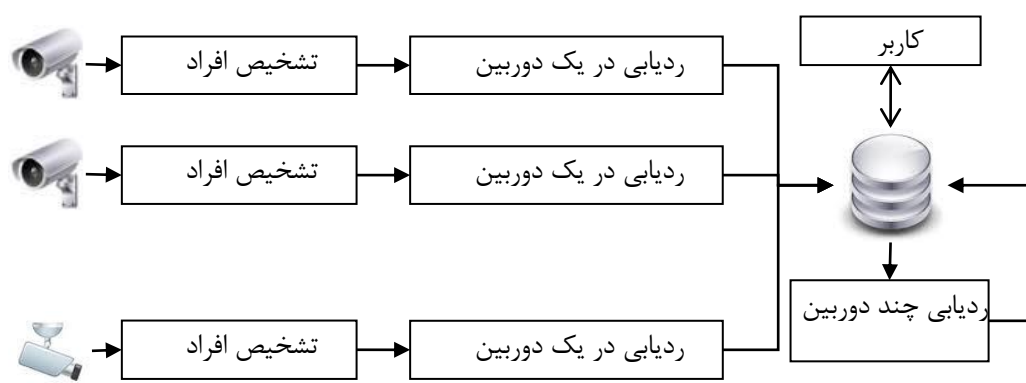
فصل ۵: پیاده سازی سیستم و نتایج حاصله

این فصل به بررسی پیاده سازی های انجام شده از مباحث و مفاهیم ارائه شده تا کنون و نتایج حاصل از آنها خواهد پرداخت. تمرکز این فصل بیشتر بر روی ویژگی تعریف شده در این پایان نامه، نتایج حاصل از آن و نحوه ی تاثیرگذاری پارامتر های آن می باشد. بدین منظور پایگاه داده ی جمع آوری شده در این پایان نامه معرفی شده و نتایج حاصل از ارائه ی الگوریتم بر روی پایگاه داده تجزیه و تحلیل خواهند شد.

در این فصل به ارائه نتایج و مشاهدات حاصل از پیاده سازی مفاهیم ارائه شده در فصل های قبلی پرداخته شده است. در ابتدای فصل توصیف کلی از سیستم پیاده سازی شده ارائه شده و در ادامه پایگاه داده ویدئویی معرفی خواهد شد. در نهایت نتایج حاصل از الگوریتم ها مورد بررسی قرار خواهد گرفت.

۵-۱- توصیف کلی سیستم پیاده سازی شده

ساختار سیستم پیشنهاد شده در **Error! Reference source not found.** نشان داده شده است. این سیستم از سه زیر سیستم مجزا تشکیل شده است. ردیابی در هر دوربین به صورت مستقل انجام شده و نتایج به دست آمده در قالب یک ساختمان داده ای استاندارد در پایگاه داده ذخیره می شوند. سپس زیر سیستم ردیابی افراد در چند دوربین با استفاده از اطلاعات موجود در پایگاه داده، ارتباط بین مشاهدات و رویدادهای اتفاق افتاده در دوربین های مختلف را برقرار می نماید.



شکل ۵-۱ ساختار سیستم ردیابی پیشنهاد شده و زیر سیستم های آن

سیستم پیشنهاد شده از یک پایگاه داده برای ذخیره ی رویدادها تشکیل شده است. رویدادهای سیستم را می توان به سه دسته ی ورود، خروج و تغییر مکان تقسیم بندی کرد. معمول ترین رویداد، حرکت فرد در محیط دید یک دوربین است. ردیابی افراد در هر یک از دوربین ها به صورت مستقل

انجام گرفته و حرکات افراد به همراه جزئیات آن^۱، در قالب یک ساختمان داده‌ی استاندارد در پایگاه داده‌ی سیستم ذخیره می‌گردند. هنگام خروج فرد از محیط دید یک دوربین، ویژگی‌های لازم برای شناسایی در مرحله‌ی بعد شناسایی شده و ذخیره می‌گردند. با ورود یک فرد به محیط دید دوربین، سیستم، ویژگی‌های به دست آمده از فرد را با رکوردهای مربوط به خروج افراد، مقایسه نموده و در صورت تطابق، اطلاعات فرد مورد نظر در پایگاه داده را بروز رسانی می‌نماید. سیستم پیاده‌سازی شده قابلیت دریافت پرس‌وجوهای^۲ مختلف را داشته و خط سیر کلی افراد در مجموعه‌ی دوربین‌ها را بر اساس پارامترهای داده شده تولید می‌نماید.

۵-۱-۱- زیر سیستم تشخیص و جدا سازی افراد

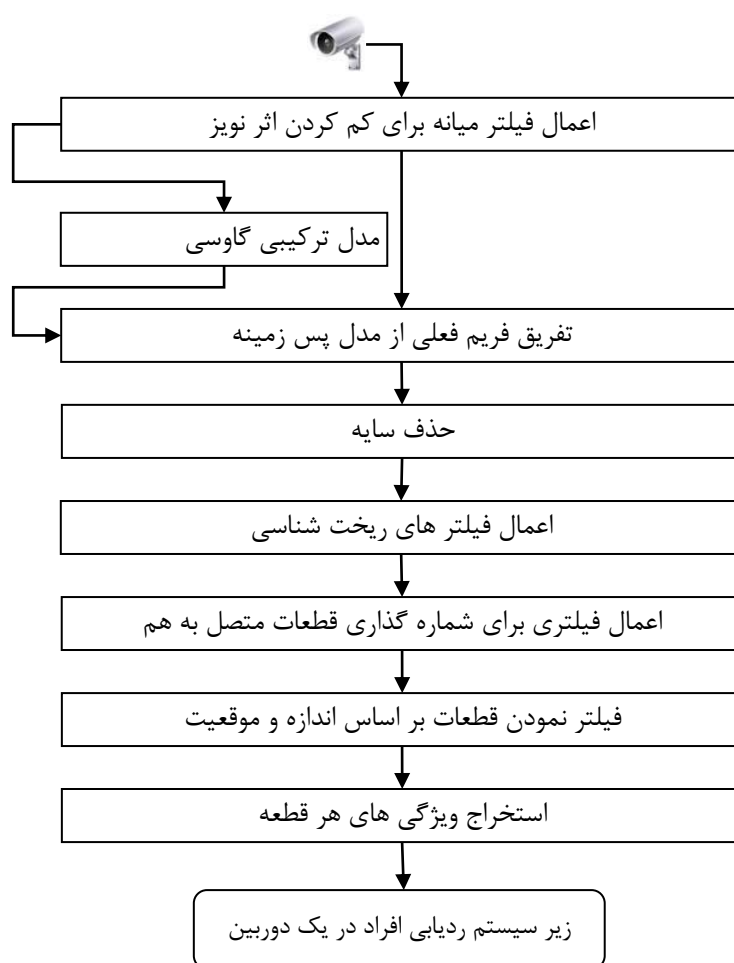
این زیر سیستم برای آشکار سازی افراد در محیط از تفاضل فریم فعلی از مدل زمینه استفاده می‌نماید. مدل زمینه با استفاده از مدل ترکیبی گاوسی ایجاد و در هر مرحله بروز رسانی می‌گردد. پس از آشکار سازی اهداف، پردازش‌هایی برای حذف نویز و سایه انجام می‌پذیرد. در مرحله‌ی بعد، مولفه‌های متصل شماره گذاری شده و سپس بر اساس اندازه فیلتر می‌شوند. اگر اندازه‌ی آنها از حد آستانه بزرگتر بودند، ویژگی‌های آنها استخراج شده و به زیر سیستم ردیابی در یک دوربین ارسال می‌گردند. شکل ۵-۲ شمای کلی این زیر سیستم را نشان می‌دهد.

۵-۱-۲- زیر سیستم ردیابی افراد در یک دوربین

این زیرسیستم با دریافت اطلاعات سطح پایین از زیر سیستم آشکار سازی افراد، با استفاده از فیلتر کالمن و ردیابی فرضیات چندگانه، افراد را در تک تک دوربین‌ها ردیابی می‌نماید. تنها اطلاعات لازم برای فیلتر کالمن در روش پیاده‌سازی شده، از اطلاعات زمانی و مکانی می‌باشند. موقعیت جسم،

^۱ جزئیات فرد شامل یک شماره‌ی منحصر به فرد، اطلاعات زمانی/ مکانی و اطلاعات ظاهری فرد می‌باشد.

^۲ Query



شکل ۵-۲ زیرسیستم آشکار سازی اهداف

سرعت در راستای افقی و عمودی و اندازه‌ی شیء ویژگی‌هایی هستند که می‌توانند در ماتریس حالت شیء تاثیر بگذارند. برای برقراری ارتباط بین فرضیات موجود با مشاهدات دریافت شده، از یک تابع هزینه به همراه الگوریتم ^۱SMM استفاده شده است. در صورتی که یک شیء جدید با یکی از فرضیات موجود تطبیق داده شد، مشخصات آن فرضیه (ماتریس حالت و پارامترهای فیلتر کالمن) بروزرسانی می‌گردد. در صورتی که یک شیء با فرضیات موجود منطبق نشود حالات متفاوتی قابل بررسی است. اولین حالت این است که واقعا یک شیء جدید وارد سیستم شده است. دومین حالت می‌تواند ناشی از چند تکه شدن تصویر یک فرد به دلیل خطا در قطعه بندی تصویر باشد و در سومین

^۱ Stable Marriage Matching

حالت ممکن است یک شی که قبلا با شی دیگری هم‌پوشانی داشته دوباره ظاهر شده باشد. برای تشخیص حالت درست از یک تابع وزن‌دهی استفاده می‌شود که با استفاده از داده‌های موجود در پایگاه داده و دانش اولیه، به هر یک از حالات وزنی را اختصاص می‌دهد. به عنوان مثال هر چه شی در مجاورت مرزهای تصویر باشد، احتمال جدید بودن شی بیشتر می‌باشد.

در هر واحد زمانی، تمام اتفاقات رخ داده در قالب ساختمان داده‌ای استاندارد در پایگاه داده ذخیره می‌گردند. ذخیره‌ی رویدادها در پایگاه داده موجب کاهش پردازش‌های لازم برای تجزیه و تحلیل‌های بعدی و همچنین کاهش حجم داده‌ی ذخیره شده می‌گردد.

هنگامی که فرد از نواحی خروجی سیستم عبور می‌نماید، سیستم ویژگی‌های مربوط به شی را استخراج نموده و در قالب رویداد خروج در پایگاه داده ذخیره می‌نماید. این اطلاعات در مرحله‌ی برقراری ارتباط بین مشاهدات در شبکه‌ی دوربین‌ها مورد استفاده قرار خواهند گرفت.

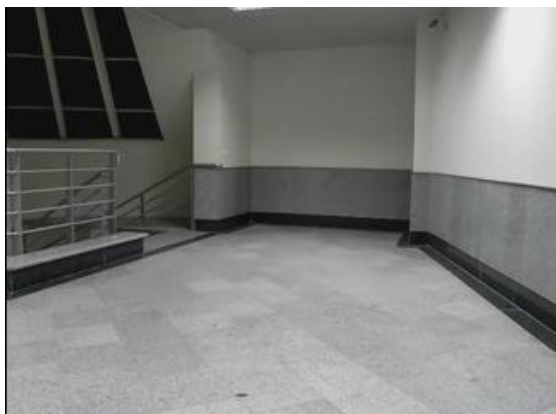
۵-۱-۳- زیر سیستم برقراری ارتباط بین مشاهدات و ردیابی در چند دوربین

هنگامی که فردی در ناحیه‌ی ورود یک دوربین قرار می‌گیرد، ویژگی معرفی شده در فصل چهارم برای تصویر آن استخراج شده و با ویژگی‌های موجود در پایگاه داده مقایسه می‌گردد. در صورتی که سیستم بتواند ویژگی جدید را با یکی از مقادیر موجود تطبیق دهد، اطلاعات فرد مربوطه شامل برچسب و شماره‌ی دوربین قبلی در پایگاه داده بروزرسانی می‌گردد. در غیراینصورت به عنوان یک شی جدید در سیستم با آن رفتار خواهد شد. معیار شباهت جنسن در این زیر سیستم برای برقراری ارتباط بین مشاهدات مورد استفاده قرار می‌گیرد.

۵-۲- مجموعه داده

به دلیل عدم وجود پایگاه داده‌ی استاندارد^۱، برای بررسی کارایی الگوریتم ارائه شده در این پایان نامه، چند مجموعه‌ی ویدئویی ضبط و گردآوری شده است. اولین مجموعه شامل پنج دوربین می باشد که پنج فرد در آن در حال رفت و آمد هستند. این ویدئوها از ساختمان مرکزی آزمایشگاه‌ها در دانشگاه صنعتی شاهرود جمع‌آوری شده است. این ساختمان از پنج طبقه تشکیل شده است که دوربین‌ها در سه طبقه‌ی آن واقع هستند. شکل ۵-۳ نواحی دید این دوربین‌ها را نمایش می‌دهد. در این مجموعه پیچیدگی‌های معمول برای یک سیستم نظارتی برای اماکن در نظر گرفته شده است. افراد در این مجموعه محدود به حرکت در یک مسیر خاص نبوده و می‌توانند در جهات مختلف در محیط حرکت کنند. الزامی به حفظ حالات حرکتی افراد وجود ندارد و افراد در بعضی از موارد با ترتیب‌های کاملاً متفاوت با خروج از یک دوربین، وارد دوربین بعدی می‌شوند. اندازه‌ی افراد در محدوده‌ی دید دوربین‌ها می‌تواند متغیر باشد. شرایط نوری در تمام محدوده‌ی بعضی از دوربین‌ها یکسان نیست. همچنین به خاطر جنس کف محدوده‌ی دید دوربین که از سنگ گرانیت می باشد، تصویر فرد تا حدی بازتاب دارد که باعث پیچیدگی در جداسازی افراد می‌گردد. سایه نیز یکی از عواملی است که بر پیچیدگی‌های صحنه می‌افزاید. در هر لحظه بیش از یک فرد در صحنه وجود دارد و در بعضی نقاط بین افراد هم پوشانی به وجود می‌آید.

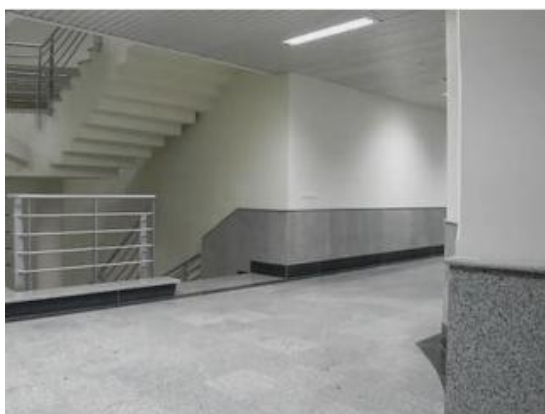
^۱ مجموعه داده‌های مشترکی برای ردیابی افراد در اینترنت وجود دارد که محققان برای مقایسه‌ی کارایی الگوریتم‌های خود می‌توانند از آنها استفاده نمایند. مجموعه تصاویر PETS و مجموعه داده‌های Etiseo از جمله‌ی مشهورترین این منابع هستند. تنها مجموعه داده‌ی موجود برای ردیابی افراد در شبکه‌ای از دوربین‌ها با دیدهای متفاوت، مجموعه-ی TERRASCOPE [۵۲] متعلق به دانشگاه کنتاکی بود که به دلیل پایان پروژه‌ی تعریف شده در این دانشگاه، دیگر در دسترس نمی‌باشد.



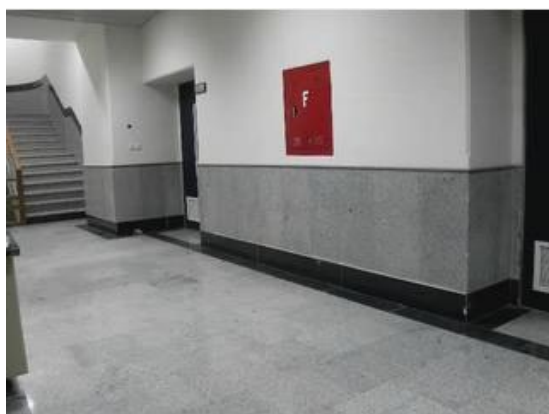
(ب)



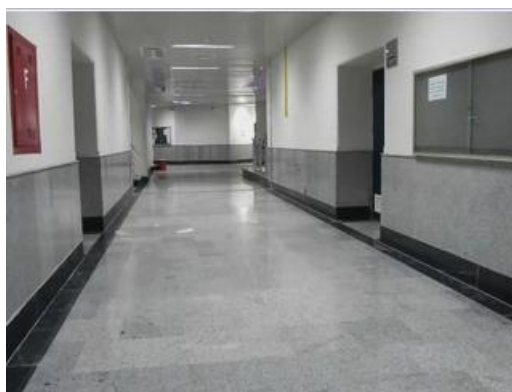
(الف)



(ت)



(پ)



(ث)

شکل ۳-۵: نمای دوربین های نصب شده در ساختمان برای جمع آوری پایگاه داده (الف) دوربین ۱، طبقه ۴ (ب) دوربین ۲ طبقه ۴ (پ) دوربین ۳، طبقه ۳ (ت) دوربین ۴، طبقه ۳ (ث) دوربین ۵، طبقه ۲

در شکل ۵-۶ تصویر به دست آمده از پنج نفر در پنج دوربین مختلف در این مجموعه داده نمایش داده شده است. داده‌های موجود در این پایگاه داده از چند جهت دارای پیچیدگی می‌باشند. همانگونه که در تصاویر قابل مشاهده می‌باشد، تصویر مربوط به افراد دارای ابعاد متفاوتی می‌باشند. بدین معنی که برخی از آنها دور از دوربین و برخی دیگر در نزدیکی دوربین ضبط شده‌اند. تصاویر از زوایای مختلفی ضبط شده‌اند. برخی از تصاویر رو به دوربین و برخی دیگر پشت به دوربین می‌باشند. به این معنی که افراد می‌توانند از جهات مختلف وارد محیط دید دوربین شوند. افراد می‌توانند در برخی از دوربین‌ها حضور نداشته باشند. به عنوان مثال فرد ۵ در دوربین ۱ حاضر نشده است. پیچیدگی دیگری که در پایگاه داده وجود دارد، عوامل مخدوش کننده‌ی تصویر یک شخص می‌باشند. سایه‌ی اشخاص تا حد زیادی توسط الگوریتم ارائه شده در فصل دوم رفع شده است، با این وجود در برخی از تصاویر، سایه تا حدی وجود دارد (تصویر فرد ۵ دوربین ۳). همچنین در برخی از تصاویر (مانند تصویر فرد ۱ در دوربین ۴)، بخش‌هایی از زمینه به عنوان تصویر فرد تشخیص داده شده است. برخی از تصاویر موجود در پایگاه داده نیز به صورت ناقص تشخیص داده شده‌اند. تشخیص نادرست افراد می‌تواند به دلیل چند تکه شدن تصویر فرد (تصویر فرد ۲ در دوربین ۱) و یا کنتراست پایین تصویر فرد و زمینه (تصویر فرد ۳ در دوربین ۲) باشد. همچنین همانگونه که به وضوح قابل مشاهده می‌باشد، تصاویر در شرایط نوری متفاوت جمع‌آوری شده‌اند.

مجموعه‌ی دوم شامل سه دوربین در آزمایشگاه‌های سیستم عامل و مدار الکترونیکی است که دو دوربین در آزمایشگاه سیستم عامل و یک دوربین دیگر در آزمایشگاه مدار الکترونیکی قرار دارد. این مجموعه نیز شامل پنج فرد است. این مجموعه شامل پیچیدگی‌های کمتری از لحاظ اندازه‌ی افراد می‌باشد؛ در مقابل هم پوشانی بین افراد در این مجموعه بیشتر از مجموعه‌ی اول می‌باشد.

جدول ۵-۱ تصاویر به دست آمده از افراد در هر یک از دوربین ها

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

استفاده از الگوریتم معرفی شده در این پایان نامه منجر به ۷۹٪ تشخیص صحیح در مجموعه‌ی اول (در مقابل ۵۱٪ برای هیستوگرام رنگ) و ۹۲٪ تشخیص صحیح در مجموعه‌ی دوم (در مقابل ۶۷٪ برای هیستوگرام رنگ) گردید. برای محاسبه‌ی دقت الگوریتم، بدترین حالت ممکن در نظر گرفته شده است. به این معنی که فرض شده است تمام دوربین‌ها با هم در ارتباط بوده و افراد پس از خروج از یک دوربین می‌توانند در فضای دید هر دوربین دیگری قرار گیرند. میزان شباهت با استفاده از مقایسه‌ی تمام تصاویر موجود در پایگاه داده با تصویر فعلی محاسبه شده و در یک جدول ذخیره می‌گردد. عدد به دست آمده برای تمام نمونه تصاویرهای مربوط به یک فرد، باید از عدد به دست آمده برای تصاویر مربوط به سایر افراد کوچکتر باشد. اگر تنها در یک مورد عدد به دست آمده برای نمونه‌های فرد دیگر، بیشتر از مقدار عدد به دست آمده برای نمونه‌های مربوط به فرد باشد، خطا برای آن مورد به حساب خواهد آمد. همانگونه که ذکر شد، این خطا بدترین حالت ممکن می‌باشد و در عمل بسیاری از ترکیب‌ها به دلیل توپولوژی سیستم و یا تشخیص‌های قبلی امکان‌پذیر نمی‌باشند. اما برای بررسی نتایج ممکن در شرایط سختگیرانه و بدون محدودیت، دقت الگوریتم با استفاده از روش ذکر شده محاسبه شده است. نحوه‌ی محاسبه برای یکی از آزمایشات در جداول ۵-۱ تا ۵-۳ نمایش داده شده است. در این جداول، مقادیری که با رنگ قرمز مشخص شده‌اند، نشان‌دهنده‌ی مقایسه‌ی تصاویر مربوط به یک فرد می‌باشند. مقادیر سفید، مربوط به نتیجه‌ی مقایسه‌ی تصویر یک فرد با تصاویر سایر افراد می‌باشند که به درستی تشخیص داده شده‌اند. خانه‌های زرد رنگ نشان‌دهنده‌ی شرایطی می‌باشند که تصویر فرد A با تصویر فرد دیگری مقایسه شده و تفاوت کمتری از حداقل تفاوت بین تصاویر مربوط به فرد A داشته و بنابراین یک تشخیص غلط می‌باشد. علامت اختصاری p نشان‌دهنده‌ی فرد و علامت اختصاری c نشان‌دهنده‌ی دوربین در جداول زیر می‌باشد.

جدول ۲-۵ مقایسه‌ی تصاویر مربوط به فرد ۲ با تصاویر موجود در پایگاه داده و نتایج حاصله

p۲-c۵	p۲-c۴	p۲-c۳	p۲-c۲	p۲-c۱	p۱-c۵	p۱-c۴	p۱-c۳	p۱-c۲	p۱-c۱	
۰.۰۵۷	۰.۰۸۰	۰.۰۶۲	۰.۰۷۰	۰.۱۱۲	۰.۰۰۱	۰.۰۰۲	۰.۰۰۱	۰.۰۰۵	۰.۰۰۰	p۱-c۱
۰.۰۴۳	۰.۰۶۲	۰.۰۴۸	۰.۰۵۶	۰.۰۹۵	۰.۰۰۳	۰.۰۰۳	۰.۰۰۶	۰.۰۰۰	۰.۰۰۵	p۱-c۲
۰.۰۵۹	۰.۰۸۱	۰.۰۶۳	۰.۰۷۱	۰.۱۱۵	۰.۰۰۳	۰.۰۰۱	۰.۰۰۰	۰.۰۰۶	۰.۰۰۱	p۱-c۳
۰.۰۵۱	۰.۰۷۲	۰.۰۵۵	۰.۰۶۴	۰.۱۰۶	۰.۰۰۲	۰.۰۰۰	۰.۰۰۱	۰.۰۰۳	۰.۰۰۲	p۱-c۴
۰.۰۵۰	۰.۰۷۱	۰.۰۵۶	۰.۰۶۳	۰.۱۰۳	۰.۰۰۰	۰.۰۰۲	۰.۰۰۳	۰.۰۰۳	۰.۰۰۱	p۱-c۵
۰.۰۱۷	۰.۰۰۹	۰.۰۱۷	۰.۰۱۴	۰.۰۰۰	۰.۱۰۳	۰.۱۰۶	۰.۱۱۵	۰.۰۹۵	۰.۱۱۲	p۲-c۱
۰.۰۰۷	۰.۰۰۹	۰.۰۱۰	۰.۰۰۰	۰.۰۱۴	۰.۰۶۳	۰.۰۶۴	۰.۰۷۱	۰.۰۵۶	۰.۰۷۰	p۲-c۲
۰.۰۰۶	۰.۰۰۴	۰.۰۰۰	۰.۰۱۰	۰.۰۱۷	۰.۰۵۶	۰.۰۵۵	۰.۰۶۳	۰.۰۴۸	۰.۰۶۲	p۲-c۳
۰.۰۰۸	۰.۰۰۰	۰.۰۰۴	۰.۰۰۹	۰.۰۰۹	۰.۰۷۱	۰.۰۷۲	۰.۰۸۱	۰.۰۶۲	۰.۰۸۰	p۲-c۴
۰.۰۰۰	۰.۰۰۸	۰.۰۰۶	۰.۰۰۷	۰.۰۱۷	۰.۰۵۰	۰.۰۵۱	۰.۰۵۹	۰.۰۴۳	۰.۰۵۷	p۲-c۵
۰.۰۵۰	۰.۰۳۷	۰.۰۴۵	۰.۰۴۴	۰.۰۵۱	۰.۰۷۰	۰.۰۷۹	۰.۰۸۴	۰.۰۷۰	۰.۰۷۷	p۳-c۱
۰.۰۳۹	۰.۰۳۷	۰.۰۳۵	۰.۰۳۹	۰.۰۵۵	۰.۰۴۰	۰.۰۴۷	۰.۰۵۰	۰.۰۴۲	۰.۰۴۵	p۳-c۲
۰.۰۴۵	۰.۰۳۸	۰.۰۳۹	۰.۰۴۳	۰.۰۵۵	۰.۰۵۶	۰.۰۶۴	۰.۰۶۷	۰.۰۵۸	۰.۰۶۲	p۳-c۳
۰.۰۵۲	۰.۰۴۰	۰.۰۵۰	۰.۰۴۵	۰.۰۴۷	۰.۰۸۳	۰.۰۹۲	۰.۰۹۷	۰.۰۸۴	۰.۰۹۱	p۳-c۴
۰.۰۶۵	۰.۰۴۶	۰.۰۵۸	۰.۰۵۶	۰.۰۵۲	۰.۱۰۰	۰.۱۱۰	۰.۱۱۵	۰.۱۰۱	۰.۱۰۸	p۳-c۵
۰.۰۲۹	۰.۰۳۶	۰.۰۳۵	۰.۰۳۳	۰.۰۵۷	۰.۰۲۴	۰.۰۲۹	۰.۰۳۳	۰.۰۲۲	۰.۰۳۰	p۴-c۱
۰.۰۳۲	۰.۰۴۲	۰.۰۳۶	۰.۰۳۷	۰.۰۶۹	۰.۰۱۳	۰.۰۱۶	۰.۰۲۱	۰.۰۰۹	۰.۰۱۹	p۴-c۲
۰.۰۳۳	۰.۰۴۵	۰.۰۳۵	۰.۰۴۲	۰.۰۷۱	۰.۰۰۸	۰.۰۱۱	۰.۰۱۳	۰.۰۰۹	۰.۰۱۱	p۴-c۳
۰.۰۳۱	۰.۰۳۷	۰.۰۳۳	۰.۰۳۵	۰.۰۶۴	۰.۰۱۶	۰.۰۱۹	۰.۰۲۴	۰.۰۱۳	۰.۰۲۲	p۴-c۴
۰.۰۳۷	۰.۰۵۰	۰.۰۴۲	۰.۰۴۶	۰.۰۷۹	۰.۰۰۸	۰.۰۱۱	۰.۰۱۵	۰.۰۰۵	۰.۰۱۳	p۴-c۵
۰.۰۴۳	۰.۰۵۸	۰.۰۴۸	۰.۰۵۲	۰.۰۸۸	۰.۰۰۷	۰.۰۱۰	۰.۰۱۲	۰.۰۰۹	۰.۰۱۰	p۵-c۲
۰.۰۴۳	۰.۰۵۷	۰.۰۴۷	۰.۰۵۰	۰.۰۸۶	۰.۰۰۹	۰.۰۱۲	۰.۰۱۴	۰.۰۱۱	۰.۰۱۲	p۵-c۳
۰.۰۳۷	۰.۰۵۰	۰.۰۳۹	۰.۰۴۶	۰.۰۷۸	۰.۰۰۸	۰.۰۱۰	۰.۰۱۳	۰.۰۰۹	۰.۰۱۱	p۵-c۴
۰.۰۴۴	۰.۰۶۲	۰.۰۵۱	۰.۰۵۵	۰.۰۹۱	۰.۰۰۵	۰.۰۰۹	۰.۰۱۱	۰.۰۰۷	۰.۰۰۸	p۵-c۵

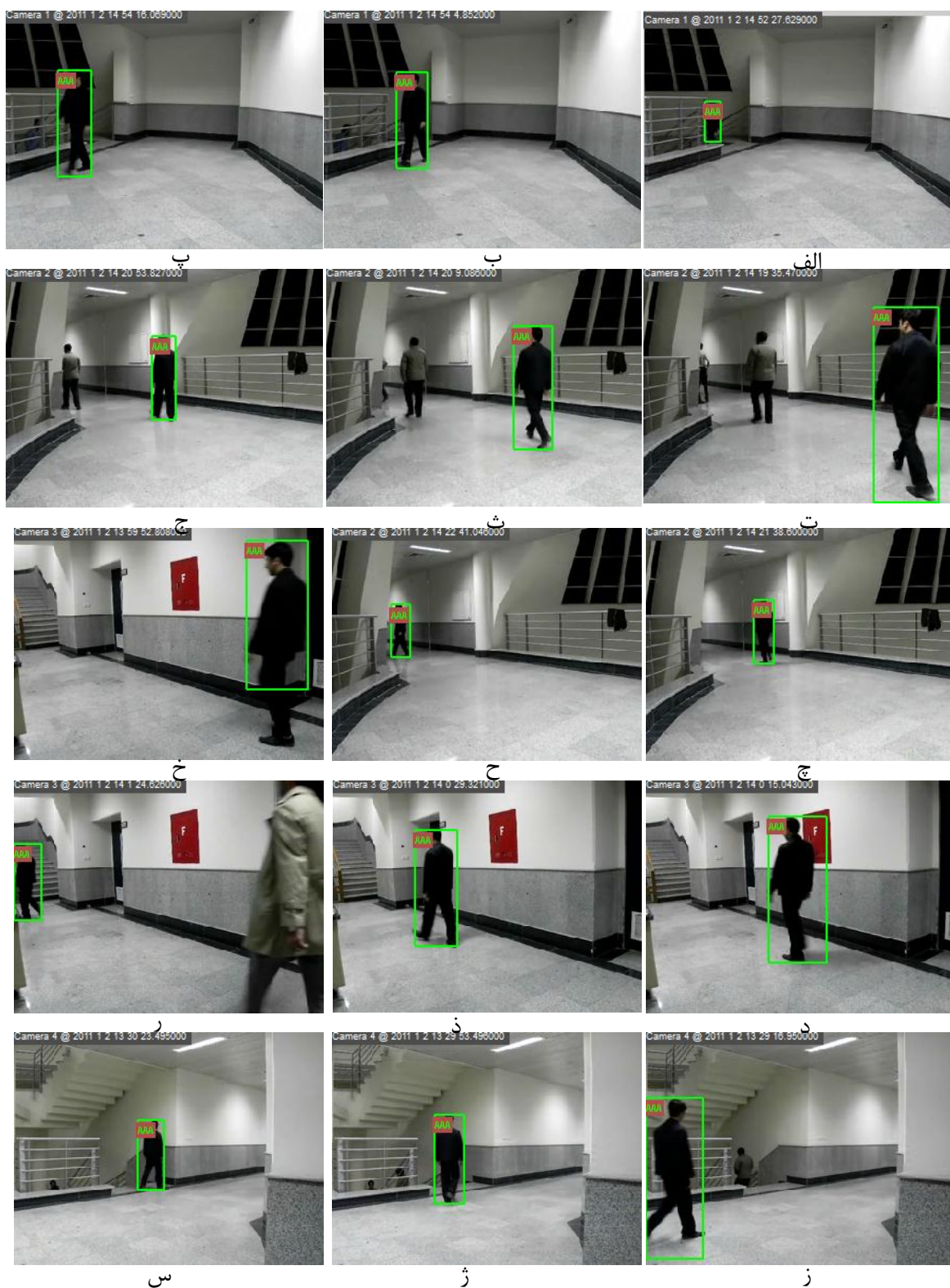
جدول ۵-۳ مقایسه‌ی تصاویر مربوط به فرد ۴۳ با تصاویر موجود در پایگاه داده و نتایج حاصله

p۴-۰۵	p۴-۰۴	p۴-۰۳	p۴-۰۲	p۴-۰۱	p۳-۰۵	p۳-۰۴	p۳-۰۳	p۳-۰۲	p۳-۰۱	
۰.۰۱۳	۰.۰۲۲	۰.۰۱۱	۰.۰۱۹	۰.۰۳۰	۰.۱۰۸	۰.۰۹۱	۰.۰۶۲	۰.۰۴۵	۰.۰۷۷	p۱-۰۱
۰.۰۰۵	۰.۰۱۳	۰.۰۰۹	۰.۰۰۹	۰.۰۲۲	۰.۱۰۱	۰.۰۸۴	۰.۰۵۸	۰.۰۴۲	۰.۰۷۰	p۱-۰۲
۰.۰۱۵	۰.۰۲۴	۰.۰۱۳	۰.۰۲۱	۰.۰۳۳	۰.۱۱۵	۰.۰۹۷	۰.۰۶۷	۰.۰۵۰	۰.۰۸۴	p۱-۰۳
۰.۰۱۱	۰.۰۱۹	۰.۰۱۱	۰.۰۱۶	۰.۰۲۹	۰.۱۱۰	۰.۰۹۲	۰.۰۶۴	۰.۰۴۷	۰.۰۷۹	p۱-۰۴
۰.۰۰۸	۰.۰۱۶	۰.۰۰۸	۰.۰۱۳	۰.۰۲۴	۰.۱۰۰	۰.۰۸۳	۰.۰۵۶	۰.۰۴۰	۰.۰۷۰	p۱-۰۵
۰.۰۷۹	۰.۰۶۴	۰.۰۷۱	۰.۰۶۹	۰.۰۵۷	۰.۰۵۲	۰.۰۴۷	۰.۰۵۵	۰.۰۵۵	۰.۰۵۱	p۲-۰۱
۰.۰۴۶	۰.۰۳۵	۰.۰۴۲	۰.۰۳۷	۰.۰۳۳	۰.۰۵۶	۰.۰۴۵	۰.۰۴۳	۰.۰۳۹	۰.۰۴۴	p۲-۰۲
۰.۰۴۲	۰.۰۳۳	۰.۰۳۵	۰.۰۳۶	۰.۰۳۵	۰.۰۵۸	۰.۰۵۰	۰.۰۳۹	۰.۰۳۵	۰.۰۴۵	p۲-۰۳
۰.۰۵۰	۰.۰۳۷	۰.۰۴۵	۰.۰۴۲	۰.۰۳۶	۰.۰۴۶	۰.۰۴۰	۰.۰۳۸	۰.۰۳۷	۰.۰۳۷	p۲-۰۴
۰.۰۳۷	۰.۰۳۱	۰.۰۳۳	۰.۰۳۲	۰.۰۲۹	۰.۰۶۵	۰.۰۵۲	۰.۰۴۵	۰.۰۳۹	۰.۰۵۰	p۲-۰۵
۰.۰۵۰	۰.۰۳۶	۰.۰۴۱	۰.۰۴۲	۰.۰۳۵	۰.۰۰۷	۰.۰۰۹	۰.۰۰۶	۰.۰۱۰	۰.۰۰۰	p۳-۰۱
۰.۰۲۹	۰.۰۲۲	۰.۰۱۷	۰.۰۲۶	۰.۰۲۰	۰.۰۲۲	۰.۰۱۴	۰.۰۰۳	۰.۰۰۰	۰.۰۱۰	p۳-۰۲
۰.۰۴۲	۰.۰۳۲	۰.۰۲۹	۰.۰۳۷	۰.۰۳۱	۰.۰۱۷	۰.۰۱۴	۰.۰۰۰	۰.۰۰۳	۰.۰۰۶	p۳-۰۳
۰.۰۵۹	۰.۰۴۴	۰.۰۴۸	۰.۰۵۱	۰.۰۳۳	۰.۰۰۹	۰.۰۰۰	۰.۰۱۴	۰.۰۱۴	۰.۰۰۹	p۳-۰۴
۰.۰۷۶	۰.۰۵۹	۰.۰۶۳	۰.۰۶۷	۰.۰۵۲	۰.۰۰۰	۰.۰۰۹	۰.۰۱۷	۰.۰۲۲	۰.۰۰۷	p۳-۰۵
۰.۰۰۹	۰.۰۰۶	۰.۰۱۱	۰.۰۰۷	۰.۰۰۰	۰.۰۵۲	۰.۰۳۳	۰.۰۳۱	۰.۰۲۰	۰.۰۳۵	p۴-۰۱
۰.۰۰۲	۰.۰۰۲	۰.۰۰۸	۰.۰۰۰	۰.۰۰۷	۰.۰۶۷	۰.۰۵۱	۰.۰۳۷	۰.۰۲۶	۰.۰۴۲	p۴-۰۲
۰.۰۰۶	۰.۰۰۸	۰.۰۰۰	۰.۰۰۸	۰.۰۱۱	۰.۰۶۳	۰.۰۴۸	۰.۰۲۹	۰.۰۱۷	۰.۰۴۱	p۴-۰۳
۰.۰۰۴	۰.۰۰۰	۰.۰۰۸	۰.۰۰۲	۰.۰۰۶	۰.۰۵۹	۰.۰۴۴	۰.۰۳۲	۰.۰۲۲	۰.۰۳۶	p۴-۰۴
۰.۰۰۰	۰.۰۰۴	۰.۰۰۶	۰.۰۰۲	۰.۰۰۹	۰.۰۷۶	۰.۰۵۹	۰.۰۴۲	۰.۰۲۹	۰.۰۵۰	p۴-۰۵
۰.۰۱۰	۰.۰۱۱	۰.۰۰۶	۰.۰۱۰	۰.۰۲۰	۰.۰۷۶	۰.۰۶۱	۰.۰۳۷	۰.۰۲۵	۰.۰۴۸	p۵-۰۲
۰.۰۱۰	۰.۰۱۰	۰.۰۰۷	۰.۰۱۰	۰.۰۲۰	۰.۰۷۴	۰.۰۶۰	۰.۰۳۷	۰.۰۲۵	۰.۰۴۶	p۵-۰۳
۰.۰۰۸	۰.۰۰۹	۰.۰۰۲	۰.۰۰۹	۰.۰۱۴	۰.۰۶۸	۰.۰۵۳	۰.۰۳۱	۰.۰۱۹	۰.۰۴۴	p۵-۰۴
۰.۰۰۷	۰.۰۱۱	۰.۰۰۴	۰.۰۰۹	۰.۰۱۷	۰.۰۷۸	۰.۰۶۲	۰.۰۳۹	۰.۰۲۵	۰.۰۵۱	p۵-۰۵

جدول ۴-۵ مقایسه‌ی تصاویر مربوط به فرد ۵ با تصاویر موجود در پایگاه داده و نتایج حاصله

p۴-۵	p۴-۴	p۴-۳	p۴-۲	
۰.۰۰۸	۰.۰۱۱	۰.۰۱۲	۰.۰۱۰	p۱-۱
۰.۰۰۷	۰.۰۰۹	۰.۰۱۱	۰.۰۰۹	p۱-۲
۰.۰۱۱	۰.۰۱۳	۰.۰۱۴	۰.۰۱۲	p۱-۳
۰.۰۰۹	۰.۰۱۰	۰.۰۱۲	۰.۰۱۰	p۱-۴
۰.۰۰۵	۰.۰۰۸	۰.۰۰۹	۰.۰۰۷	p۱-۵
۰.۰۹۱	۰.۰۷۸	۰.۰۸۶	۰.۰۸۸	p۲-۱
۰.۰۵۵	۰.۰۴۶	۰.۰۵۰	۰.۰۵۲	p۲-۲
۰.۰۵۱	۰.۰۳۹	۰.۰۴۷	۰.۰۴۸	p۲-۳
۰.۰۶۲	۰.۰۵۰	۰.۰۵۷	۰.۰۵۸	p۲-۴
۰.۰۴۴	۰.۰۳۷	۰.۰۴۳	۰.۰۴۳	p۲-۵
۰.۰۵۱	۰.۰۴۴	۰.۰۴۶	۰.۰۴۸	p۳-۱
۰.۰۲۵	۰.۰۱۹	۰.۰۲۵	۰.۰۲۵	p۳-۲
۰.۰۳۹	۰.۰۳۱	۰.۰۳۷	۰.۰۳۷	p۳-۳
۰.۰۶۲	۰.۰۵۳	۰.۰۶۰	۰.۰۶۱	p۳-۴
۰.۰۷۸	۰.۰۶۸	۰.۰۷۴	۰.۰۷۶	p۳-۵
۰.۰۱۷	۰.۰۱۴	۰.۰۲۰	۰.۰۲۰	p۴-۱
۰.۰۰۹	۰.۰۰۹	۰.۰۱۰	۰.۰۱۰	p۴-۲
۰.۰۰۴	۰.۰۰۲	۰.۰۰۷	۰.۰۰۶	p۴-۳
۰.۰۱۱	۰.۰۰۹	۰.۰۱۰	۰.۰۱۱	p۴-۴
۰.۰۰۷	۰.۰۰۸	۰.۰۱۰	۰.۰۱۰	p۴-۵
۰.۰۰۲	۰.۰۰۲	۰.۰۰۱	۰.۰۰۰	p۵-۲
۰.۰۰۴	۰.۰۰۴	۰.۰۰۰	۰.۰۰۱	p۵-۳
۰.۰۰۲	۰.۰۰۰	۰.۰۰۴	۰.۰۰۲	p۵-۴
۰.۰۰۰	۰.۰۰۲	۰.۰۰۴	۰.۰۰۲	p۵-۵

شکل ۵-۷ مجموعه ای از نتیجه ی به دست آمده برای ردیابی یک فرد را نمایش می دهد.



شکل ۵-۴ نمونه ای از ردیابی یک فرد در مجموعه ی دوربین ها از شماره ۱ تا شماره ۴، شروع از راست به چپ

۵-۳- بررسی پارامترهای تاثیرگذار

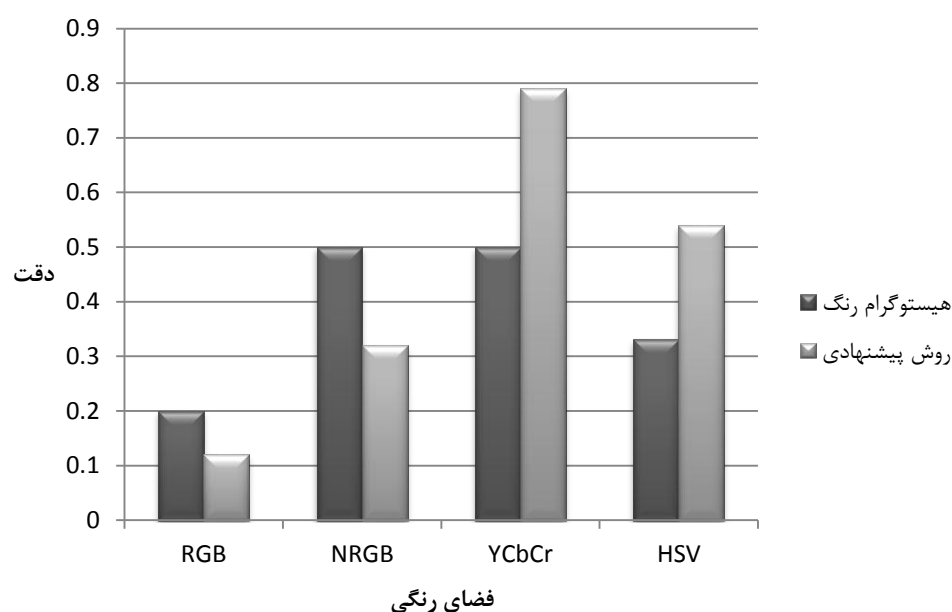
در این بخش تاثیر پارامترهای مختلف بر روی نتیجه‌ی نهایی کسب شده مورد بررسی قرار می‌گیرد. پارامترهای زیادی در برقراری ارتباط بین مشاهدات تاثیر دارند که مهمترین آنها عبارتند از

- فضای رنگی انتخاب شده
- تعداد عناصر موجود در بردار ویژگی
- معیار شباهت دو بردار
- تعداد قطر ها در استخراج ویژگی از ماتریس هم‌جواری
- تجمعی بودن توزیع
- نحوه ی ترکیب نتایج حاصل از هر مولفه (در فضاهای رنگی سه بعدی)

برای بررسی کارایی الگوریتم ارائه شده و مشاهده‌ی تاثیر پارامترهای مختلف، ترکیب‌های موجود برای پارامترهای ذکر شده مورد بررسی قرار گرفت. در این بخش برخی از نتایج به دست آمده مورد تجزیه و تحلیل قرار خواهند گرفت. به دلیل تعدّد حالات ممکن، در هر قسمت برخی از پارامترهای ثابت فرض شده و تاثیر سایر پارامترها مورد بررسی قرار گرفته است

• مقایسه‌ی کارایی ویژگی معرفی شده (بردار قطری ماتریس هم‌رخداد) با رنگ

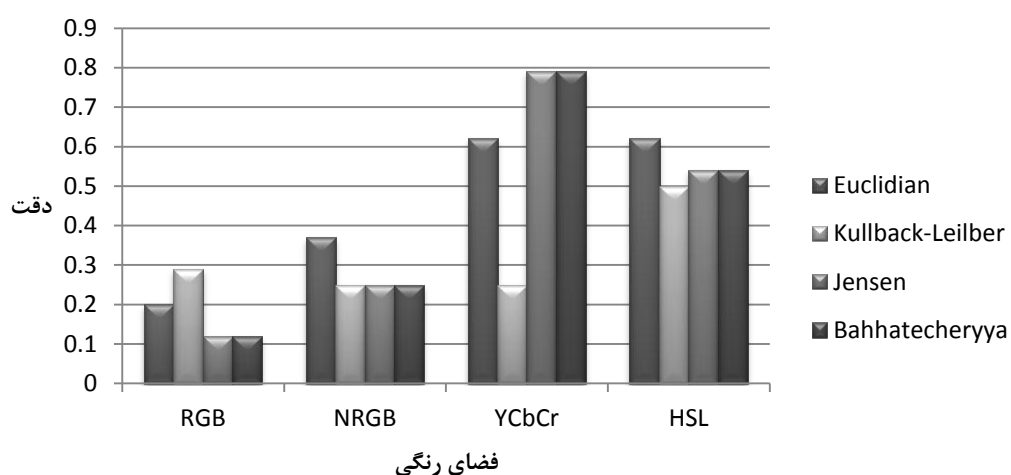
اولین بررسی مربوط به اثبات کارایی ویژگی ارائه شده نسبت به هیستوگرام رنگ می باشد. برای این منظور، تطبیق مشاهدات مربوط به افراد، توسط هیستوگرام رنگ و ویژگی پیشنهاد شده (بردار قطری ماتریس هم‌رخداد) در فضاهای رنگی مختلف، مورد بررسی قرار گرفته‌اند. معیار بررسی شباهت تابع جنسن انتخاب شده است. همانگونه که در شکل ۵-۵ مشاهده می نمایید، دقت ویژگی ارائه شده در فضاهای رنگی که اطلاعات روشنایی را به صورت جدا ذخیره می نمایند، بالاتر است.



شکل ۵-۵ دقت هیستوگرام رنگ و ویژگی پیشنهاد شده در فضاهای رنگی مختلف

• بررسی معیار شباهت

عامل بسیار مهم دیگری که می‌تواند بر کارایی سیستم تاثیر بسزایی داشته باشد، انتخاب معیار مقایسه می باشد. معیارهای فاصله‌ی اقلیدسی، تابع جنسن، ضریب باتاچریا و ضریب واگرایی کولبک-لیلبر بر روی داده‌ها آزمایش شده‌اند که نتیجه این مقایسه را در شکل ۵-۶ نشان داده شده است.

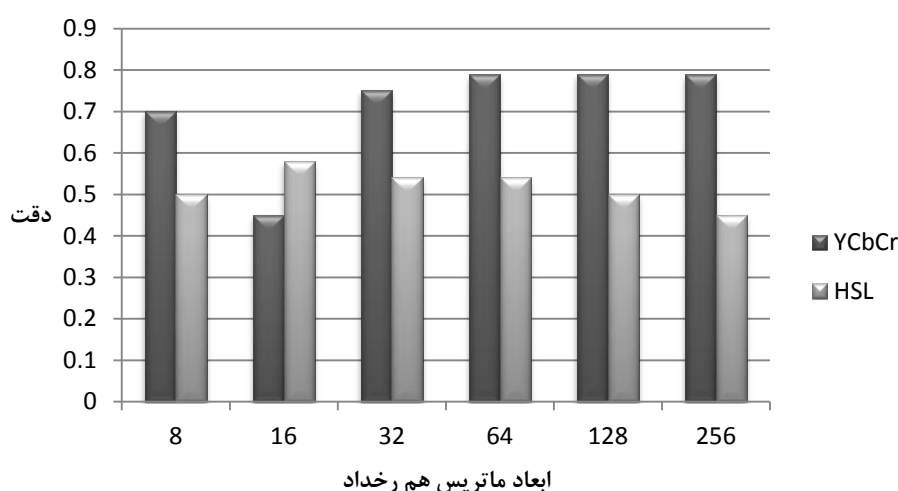


شکل ۵-۶: نتایج به دست آمده استفاده از معیارهای شباهت بر روی ویژگی تعریف شده

از شکل ۵-۶ می‌توان دریافت، معیار جنسن و ضریب باتاچریا در فضای رنگ تعریف شده منجر به نتایج بهتری شده‌اند.

• تاثیر ابعاد ماتریس هم‌رخداد

ابعاد ماتریس هم‌رخداد می‌تواند تاثیر بسزایی در دقت الگوریتم داشته باشد. ابعاد کمتر به معنی این است که هر مولفه، بازه‌ی گسترده‌تری از رنگ‌ها را شامل می‌شود. در عمل تعداد مولفه‌های کمتر به دلیل حذف جزئیات مربوط به رنگ‌ها، باعث کاهش کارایی می‌گردد. از طرفی تعداد مولفه‌های بیشتر، بار محاسباتی بیشتری را به سیستم تحمیل می‌نماید و امکان بروز خطا در الگوریتم مشابهت به دلیل اثر نویز یا تغییرات جزئی نور بیشتر می‌گردد. نتایج به دست آمده در شکل ۵-۱۰ نمایش داده شده است. همانگونه که مشاهده می‌شود فضای رنگ YCbCr منجر به نتایج بهتری شده است.

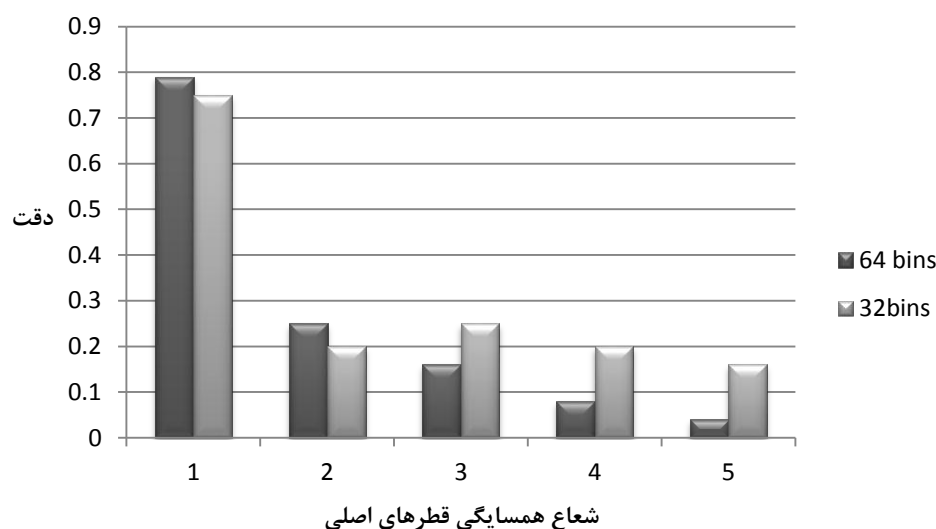


شکل ۵-۷ تاثیر ابعاد ماتریس هم‌جواری در کارایی سیستم

• تاثیر تعداد قطر‌ها

همانگونه که در فصل قبل ذکر شد، برای تصاویر افراد، تجمع مقادیر غیر صفر ماتریس هم‌رخداد نزدیک قطر اصلی می‌باشد. برای بررسی تعداد قطر‌های لازم، مقادیر مختلف از یک تا ۹ قطر (شعاع ۱

تا ۵) بر روی ماتریس هم رخداد بررسی شد که نتیجه ی آن را در شکل ۵-۱۱ مشاهده می نمایید. نتیجه ی این بررسی بسیار جالب است، با افزایش تعداد قطرها دقت الگوریتم به شدت کاهش می یابد. دلیل این امر، یکنواختی تقریبی رنگ افراد است. همانگونه که مشاهده می گردد، شیب افت کیفیت با پایین آمدن ابعاد ماتریس کمتر می شود.



شکل ۵-۸ تاثیر تعداد قطرهای اصلی بر کارایی سیستم

۵-۴- نتیجه گیری

در این فصل نتایج استفاده از ویژگی معرفی شده در این پایان نامه، برای برقراری ارتباط بین مشاهدات به دست آمده در مجموعه ی دوربین ها مورد بررسی قرار گرفته است. ابتدا شمای کلی از سیستم موجود معرفی شده و به دنبال آن مجموعه داده های ویدئویی موجود در پایگاه داده به همراه پیچیدگی های ممکن در آن توضیح داده شده است. برای اثبات کارایی ابتدا نتایج به دست آمده با هیستوگرام رنگ در فضا های رنگی مختلف مقایسه شده و سپس پارامترهای تاثیرگذار بر کارایی ویژگی مورد بررسی قرار گرفته است.

فصل ۶: نتیجه گیری و پیشنهادات برای کارهای آینده

این فصل به بررسی نتایج حاصل از پایان نامه پرداخته و در ادامه پیشنهاداتی را برای ادامه ی کار مطرح خواهد نمود.

۶-۱- نتیجه گیری

در این پایان نامه، ردیابی افراد در شبکه ای از دوربین ها با دیدهای مجزا (هیچ نقطه ی مشترکی بین دو دوربین وجود ندارد)، در محیط های بسته مانند ادارات و مراکز خرید مورد بررسی قرار گرفته است. سیستم های ردیابی افراد، علاوه بر کاربردهای امنیت، کاربرد های فراوان دیگری نیز در زمینه- های هوش فراگیر، هوش تجاری و مدیریتی دارند.

اغلب روش های موجود برای ردیابی افراد در چند دوربین از اطلاعات ظاهری افراد برای برقراری تناظر بین مشاهدات مربوط به یک فرد بهره می گیرند. رنگ و بافت دو ویژگی سطح پایین هستند که عموماً برای تشخیص اشیا از آنها استفاده می شود. رنگ هر شیء به صورت هیستوگرام و یا نمایش های مبتنی بر رنگ های غالب نشان داده می شود. استفاده از هیستوگرام، اطلاعات مکانی مقادیر رنگ را نادیده می گیرد. در مقابل استفاده از بافت، عموماً از اطلاعات رنگی شیء چشم پوشی می نماید.

در این پایان نامه یک ویژگی جدید برای مساله ی ردیابی اهداف در شبکه ای از دوربین ها با دید های مجزا برای مکان های بسته ارائه شده است. این روش به صورت همزمان از اطلاعات بافتی و همچنین اطلاعات رنگی استفاده می نماید. به این ترتیب که به جای استفاده از اطلاعات رنگ، از عناصر واقع بر روی قطر های ماتریس همجواری تصویر شیء، در فضای رنگ YCbCr استفاده شده است. این ویژگی نسبت به رنگ ثبات بیشتری داشته و از حساسیت کمتری برخوردار می باشد.

پایگاه داده ی ویدئویی تعریف شده برای ارزیابی کارایی الگوریتم ارائه شده، از چندین جهت پیچیده بوده و برخی از محدودیت هایی را که در سیستم های موجود وجود دارد، در نظر نگرفته است. تصویر مربوط به افراد دارای ابعاد بوده و در فواصل مختلف از دوربین ضبط شده اند. تصاویر از زوایای مختلفی ضبط شده اند؛ برخی از تصاویر رو به دوربین و برخی دیگر پشت به دوربین می باشند. همچنین تعداد افراد در دوربین های مختلف می تواند متفاوت بوده و افراد می توانند آزادانه و بدون رعایت قوانین خاص

برای حفظ حالات حرکت، در محیط رفت و آمد داشته باشند. پیچیدگی دیگری که در پایگاه داده وجود دارد، عوامل مخدوش کننده‌ی تصویر یک شخص از جمله وجود سایه و نویز می‌باشد. الگوریتم ارائه شده تا حدی، توانایی کار با داده های ناقص و مخدوش را نیز دارد. این داده ها عموماً به دلیل مشکلات ناشی از قطعه‌بندی نادرست می باشد. گاهی تصویر فرد شامل بخش‌هایی از زمینه به عنوان تصویر فرد می‌باشد. برخی از تصاویر نیز به صورت ناقص تشخیص داده شده‌اند. تشخیص ناقص افراد می‌تواند به دلیل چند تکه‌شدن تصویر فرد و یا کنتراست پایین تصویر فرد و زمینه باشد. به علاوه نمونه ها در شرایط نوری متفاوت جمع‌آوری شده‌اند.

سیستم ارائه شده از دو دیدگاه متفاوت در ردیابی تکی و ردیابی در چند دوربین استفاده کرده است. برای ردیابی افراد در یک دوربین تنها ویژگی های مکانی - زمانی به همراه معادلات تغییر حالت مورد استفاده قرار گرفته است. استفاده از این ویژگی به همراه فیلتر کالمن، قابلیت ردیابی افراد در موارد بروز هم‌پوشانی کامل را امکان پذیر می‌سازد. علاوه بر این استقلال زیر سیستم ردیابی افراد در یک دوربین از ظاهر افراد، به سیستم امکان ردیابی افراد هنگام تغییر پوشش در محیط را تا حدودی فراهم نموده و افراد می‌توانند تا هنگامی که ناحیه‌ی تشخیص داده شده‌ی آنها تغییر ننماید، پوشش خود را تغییر دهند. در مقابل ردیابی در چند دوربین به هیچ وجه از اطلاعات زمانی - مکانی استفاده ننموده و تنها وابسته به ویژگی های ظاهری افراد می باشد. بنابر این فرد باید دارای پوشش یکسان در لحظه ی خروج از دید یک دوربین و زمان ورود به دوربین دیگر باشد.

سیستم پیشنهاد شده به دلیل استفاده از یک معماری چند لایه توزیع شده، توانایی مدیریت حجم وسیعی از داده‌ها را دارد. تمام رویدادهای اتفاق افتاده در سیستم، در قالب یک ساختمان داده‌ی استاندارد در پایگاه داده ذخیره می‌شوند. بنابر این سیستم قادر به دریافت پرس و جوهای مختلف از کاربر و تولید خروجی بر اساس پارامتر های ورودی می باشد. پرس و جو ها می توانند اطلاعات را در قالب نوع رویداد، زمان و مکان رویداد، فرد مورد نظر و سایر پارامتر های موجود فیلتر نمایند.

فیلترینگ بر اساس زمان، مکان و فرد مورد نظر به عنوان مثال مورد ارزیابی قرار گرفته است. همچنین سیستم می تواند با دریافت پارامترهای ورودی، تمام اطلاعات موجود را در قالب یک فیلم تولید نماید.

برای ارزیابی ویژگی معرفی شده و تاثیر پارامترهای مختلف بر آن آزمایش‌های متعددی صورت پذیرفته است. پارامترهایی از قبیل فضاهای رنگی مختلف، ابعاد ماتریس هم‌رخداد، تعداد قطره‌های اصلی و معیار شباهت مورد بررسی قرار گرفته است. بهترین نتیجه هنگام استفاده از فضای رنگی YCbCr به دست آمده است. در این فضای رنگی ویژگی روشنایی به طور کامل از سایر مولفه‌ها جدا شده است، بنابراین ویژگی‌های به دست آمده در این فضای رنگ دارای ثبات بیشتری از سایر فضاهای رنگی می باشند. پس از این فضای رنگی، فضای رنگ HSL نیز تا حدی منجر به جواب‌های خوب شده است. تعداد ابعاد ماتریس هم رخداد، تاثیر به سزایی در کارایی سیستم ردیابی دارد. ابعاد این ماتریس باید به گونه‌ای انتخاب شود که قابلیت تمییز رنگ‌های افراد در سیستم وجود داشته باشد. نتایج نشان می‌دهد کمی‌سازی ابعاد ماتریس به مقادیر ۳۲، ۶۴ و ۱۲۸ دارای بیشترین کارایی در فضای رنگ مطرح شده دارند. همچنین معیار جنسن در فضای رنگ YCbCr و معیار باتاچریا در فضای رنگ HSV منجر به بهترین جواب‌ها شده اند.

نکته ی دیگر کارایی ویژگی مستقل از مقیاس (SIFT) می باشد. اگرچه این ویژگی در بسیاری از کاربردهای تشخیص اشیا منجر به نتیجه های خوبی شده است اما بنا به ماهیت روان بودن و تغییر پذیری لباس افراد استفاده از این ویژگی نتایج خوبی را در برنداشت. همچنین به دلیل وضوح و کیفیت پایین تصویر دریافتی در سیستم‌های نظارتی، عملاً استفاده از این ویژگی را برای برقراری ارتباط بین مشاهدات مربوط به یک فرد غیر ممکن می سازد. در آزمایشات انجام پذیرفته در این پایان نامه، به طور متوسط ۷ نقطه‌ی کلیدی در تصویر مربوط به هر فرد به دست آمد که نسبت نقاط

کلیدی مربوط به یک فرد، که امکان انطباق داشتند، حتی در یک دوربین، بسیار کم و نزدیک به دو نقطه‌ی کلیدی بود.

۶-۲- پیشنهادات

کارهای موجود در زمینه‌ی ردیابی افراد در چند دوربین با دیدهای مجزا، از پایگاه داده‌ی مخصوص به خود استفاده می نمایند. دلیل این امر عدم وجود پایگاه داده‌ی استاندارد در این زمینه می باشد. تنها پایگاه داده‌ای در این زمینه، پایگاه داده‌ی Terrascope بود که در حال حاضر امکان استفاده از آن وجود ندارد. به دلیل استفاده از داده‌های مختلف، -که قطعاً از نظر شرایط و میزان پیچیدگی متفاوت می باشند- مقایسه‌ی کارهای انجام پذیرفته شده را با مشکل مواجه می نماید. بنابراین نیاز به یک پایگاه داده‌ی استاندارد به شدت احساس می گردد. ایجاد پایگاه داده می تواند به عنوان اولین شاید مهمترین گام برای کارهای آینده باشد.

در این سیستم نحوه‌ی ارتباط بین محیط دوربین ها در نظر نگرفته شده است. با تغییرات اندک می توان کارایی الگوریتم موجود برای مقایسه و تطبیق ویژگی ها را افزایش داد.

اگرچه تمرکز این پروژه بیشتر بر روی ردیابی در چند دوربین بوده است، اما ردیابی در تک تک دوربین ها نیز مهم بوده و مشکلات ناشی از این مرحله، در نتیجه‌ی نهایی تاثیر بسزایی دارد. در ادامه‌ی این بخش به برخی از این مشکلات و پیشنهادی برای کارهای آینده در این زمینه اشاره خواهد شد. در مسائل واقعی افراد می توانند پوشش خود را تغییر دهند. در این پروژه به این نکته توجه شده است و با تکیه بر اطلاعات زمانی و مکانی، تا حدی این مساله رفع شده است. در سیستم ارائه شده فرد می تواند پوشش خود را به شرط عدم تغییر اساسی در ناحیه‌ی تشخیص داده شده در تصویر، تعویض نماید. تکیه تنها بر استفاده از اطلاعات مکانی و زمانی، اگرچه مزایایی دارد، اما باعث محدودیت هایی

نیز می گردد. یکی از این محدودیت ها مساله ی چند تکه شدن^۱ تصویر فرد می باشد. در این مشکل بنا به دلایل مختلف، تصویر یک فرد به چند قسمت تقسیم می گردد که باعث تشخیص های اشتباه می گردد. علاوه بر این اگر این مشکل در ناحیه ی ورودی و یا خروجی اتفاق بیفتد، ممکن است باعث تشخیص نادرست در الگوریتم ارتباط مشاهدات گردد.

هر چند هم پوشانی بین افراد تا حدی با استفاده از اطلاعات زمانی- مکانی مرتفع شده است، اما در مواردی که هم پوشانی ها تشدید می گردد - هم پوشانی بین تعداد زیادی از افراد- و یا هنگامی که افراد در حین هم پوشانی حالت حرکتی خود را تغییر دهند، این الگوریتم قادر به تشخیص نمی باشد. استفاده از سایر ویژگی ها مانند بافت و یا رنگ به شرطی که کلیات مساله را تغییر ندهد - به عنوان مثال تغییر لباس افراد- می تواند موجب بهبود کارایی الگوریتم تشخیص هم پوشانی گردد. مساله ی دیگری که در برخی موارد روال کار برنامه را تحت تاثیر قرار می دهد، تشخیص قسمتی از بدن فرد به عنوان زمینه می باشد. این مساله می تواند ناشی از کنتراست پایین بین زمینه و لباس فرد و یا تشخیص قسمتی از لباس فرد به عنوان سایه باشد. این مساله خصوصا در لباس های با رنگ روشن دو شرایط نوری کم اتفاق می افتد. برای حل این مشکل نیاز به تحقیق بیشتر بر روی روش های مدل سازی زمینه می باشد.

سیستم پیاده سازی شده از فیلتر کالمن برای ردیابی افراد در محیط استفاده نموده است. در مواردی که هدف دنبال کردن تنها یک فرد باشد، استفاده از فیلتر پارتیکل ممکن است منجر به نتایج بهتری گردد.

^۱ Fragmentation

- [1] A. Yilmaz, O. Javed, and M. Shah, "Object Tracking : A Survey," *ACM Computing surveys*, vol. 38, 2006.
- [2] O. Javed and M. Shah, *Automated Multi-Camera Surveillance Algorithms and Practice*, Springer, 2008.
- [3] C. Madden, "Tracking People Across Disjoint Camera Views," 2009.
- [4] J.L. Barron, D.J. Fleet, S.S. Beauchemin, and T.A. Burkitt, "Performance of Optical Flow techniques," *IEEE Conference on Computer Vision and Pattern Recognition.*, 1992, pp. 236-242.
- [5] S. Ribaric, G. Adrinek, and S. Segvic, "Real-time active visual tracking system," *Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference (IEEE Cat. No.04CH37521)*, Ieee, 2004, pp. 231-234.
- [6] W. Hu, T. Tan, L. Wang, and S. Maybank, "A Survey on Visual Surveillance of Object Motion and Behaviors," *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, IEEE, 2004, pp. 334-352.
- [7] Z. Tang, Z. Miao, and Y. Wan, "Background Subtraction Using Running Gaussian Average and Frame Difference," *Entertainment Computing-ICEC 2007*, 2007, p. 411-414.
- [8] J. Rymel, J. Renno, D. Greenhill, J. Orwell, and G. a Jones, "Adaptive eigen-backgrounds for object detection," *2004 International Conference on Image Processing, 2004. ICIP '04.*, 2004, pp. 1847-1850.
- [9] C. Stauffer and W.E.L. Grimson, "Adaptive background mixture models for real-time tracking," *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, IEEE, 2002, pp. 246-252.
- [10] D. Harwood and L. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, 2000, p. 809-830.
- [11] M. Piccardi, "Background subtraction techniques: a review," *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*, Ieee, 2004, pp. 3099-3104.
- [12] D.H. Parks and S.S. Fels, "Evaluation of Background Subtraction Algorithms with Post-Processing," *2008 IEEE Fifth International Conference on Advanced Video and Signal Based Surveillance*, Ieee, 2008, pp. 192-199.
- [13] A. Cavallaro, E. Salvador, T. Ebrahimi, and Q. Mary, "DETECTING SHADOWS IN IMAGE SEQUENCES," 2004, pp. 165-174.

- [14] C. Wang and W. Zhang, "A Robust Algorithm for Shadow Removal of Foreground Detection in Video Surveillance," *2009 Asia-Pacific Conference on Information Processing*, Jul. 2009, pp. 422-425.
- [15] R. Polana and R. Nelson, "Low level recognition of human motion (or how to get your man without finding his body parts)," *Proceedings of 1994 IEEE Workshop on Motion of Non-rigid and Articulated Objects*, 1994, pp. 77-82.
- [16] J. Li, C.S. Chua, and Y.K. Ho, "Color based multiple people tracking," *Control, Automation, Robotics and Vision, 2002. ICARCV 2002. 7th International Conference on*, IEEE, 2003, p. 309-314.
- [17] P. Perez, C. Hue, J. Vermaak, and M. Gangnet, "Color-based probabilistic tracking," *Computer Vision—ECCV 2002*, 2002, p. 661-675.
- [18] W. Lu and Y.P. Tan, "A color histogram based people tracking system," *Circuits and Systems, 2001. ISCAS 2001. The 2001 IEEE International Symposium on*, IEEE, 2002, p. 137-140.
- [19] C.J. Poelman and T. Kanade, "A Paraperspective Factorization Method for Shape and Motion Recovery," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, vol. 19, 1997, pp. 206-218.
- [20] Y. Chen, Y. Rui, and T.S. Huang, "JPDAF Based HMM or Real-Time Contour Tracking," *Computing*, 2001.
- [21] Y. Tsai and A. Averbuch, "Automatic segmentation of moving objects in video sequences: a region labeling approach," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 12, 2002, p. 597-612.
- [22] D.S. Jang and H.I. Choi, "Active models for tracking moving objects," *Pattern Recognition*, vol. 33, Jul. 2000, p. 1135-1146.
- [23] C. He, Y.F. Zheng, and S.C. Ahalt, "Object tracking using the Gabor wavelet transform and the golden section algorithm," *Multimedia, IEEE Transactions on*, vol. 4, Dec. 2003, p. 528-538.
- [24] C.T. Hsieh, E. Lai, Y.K. Wu, and C.K. Liang, "Robust, real time people tracking with shadow removal in open environment," *Control Conference, 2004. 5th Asian*, IEEE, 2005, p. 901-905.
- [25] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "P finder : Real-Time Tracking of the Human Body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, 1997, pp. 780-785.
- [26] D.P. Huttenlocher, J.J. Noh, and W.J. Rucklidge, "Tracking non-rigid objects in complex scenes," *1993 (4th) International Conference on Computer Vision*, IEEE Computer Society Press, 1993, pp. 93-101.

-
- [27] S. McKenna, "Tracking Groups of People," *Computer Vision and Image Understanding*, vol. 80, Oct. 2000, pp. 42-56.
 - [28] N. Peterfreund, "Robust tracking of position and velocity with Kalman snakes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, Jun. 2002, p. 564-569.
 - [29] C.E. Erdem, a M. Tekalp, and B. Sankur, "Video object tracking with feedback of performance measures," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, Apr. 2003, pp. 310-324.
 - [30] G. Welch and G. Bishop, "An introduction to the Kalman filter," *University of North Carolina at Chapel Hill, Chapel Hill, NC*, 1995.
 - [31] M.I. Ribeiro, "Kalman and extended kalman filters: Concept, derivation and properties," *Institute for Systems and Robotics, Lisboa Portugal*, 2004.
 - [32] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on signal processing*, vol. 50, 2002, p. 174-188.
 - [33] K. Nummiaro, E. Koller-Meier, and L. Van Gool, "An adaptive color-based particle filter," *Image and Vision Computing*, vol. 21, Jan. 2003, pp. 99-110.
 - [34] D. Comaniciu and V. Ramesh, "Real-time tracking of non-rigid objects using mean shift," *US Patent 6,590,999*, Jul. 2003.
 - [35] K. Jeong, "OBJECT MATCHING IN DISJOINT CAMERAS USING A COLOR TRANSFER APPROACH," 2006.
 - [36] T. D'Orazio, P.L. Mazzeo, and P. Spagnolo, "Color Brightness Transfer Function evaluation for non overlapping multi camera tracking," *Distributed Smart Cameras, 2009. ICDSC 2009. Third ACM/IEEE International Conference on*, IEEE, 2009, p. 1-6.
 - [37] C. Madden, E.D. Cheng, and M. Piccardi, "Tracking people across disjoint camera views by an illumination-tolerant appearance representation," *Machine Vision and Applications*, vol. 18, Mar. 2007, p. 233-247.
 - [38] K. Jeong and C. Jaynes, "Object matching in disjoint cameras using a color transfer approach," *Machine Vision and Applications*, vol. 19, May. 2007, pp. 443-455.
 - [39] J. Orwell, P. Remagnino, and G. a Jones, "Multi-camera colour tracking," *Proceedings Second IEEE Workshop on Visual Surveillance (VS'99) (Cat. No.98-89223)*, IEEE Comput. Soc, , pp. 14-21.

- [40] B. Prosser, S. Gong, T. Xiang, and Q. Mary, "Multi-camera matching using bi-directional cumulative brightness transfer functions," *Proceedings of the British Machine Vision Conference*, Citeseer, 2008.
- [41] T. Darrell, G. Gordon, M. Harville, and J. Woodfill, "Integrated person tracking using stereo, color, and pattern detection," *International Journal of Computer Vision*, vol. 37, 2000, p. 175–185.
- [42] 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 Rim Conference on Multimedia*, IEEE Press, 2003, p. 1335–1340.
- [43] C. BenAbdelkader, R. Cutler, and L. Davis, "Person identification using automatic height and stride estimation," *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, IEEE, 2005, p. 377–380.
- [44] V. Kettner, "Bayesian multi-camera surveillance," *Computer Vision and Pattern*, vol. 00, 2002, pp. 253-259.
- [45] a Rahimi, B. Dunagan, and T. Darrell, "Simultaneous calibration and tracking with a network of non-overlapping sensors," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, pp. 187-194.
- [46] O. Javed and K. Shafique, "Appearance Modeling for Tracking in Multiple Non-Overlapping Cameras," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 26-33.
- [47] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, Nov. 2004, p. 91–110.
- [48] D.R. Kisku, A. Rattani, E. Grosso, and M. Tistarelli, "Face Identification by SIFT-based Complete Graph Topology," *2007 IEEE Workshop on Automatic Identification Advanced Technologies*, Jun. 2007, pp. 63-68.
- [49] R.M. Haralick, K. Shanmugam, and I.H. Dinstein, "Textural features for image classification," *IEEE Transactions on systems, man and cybernetics*, vol. 3, Nov. 1973, p. 610–621.
- [50] R.C. Gonzalez and R.E. Woods, *Digital Image Processing (3rd Edition)*, Prentice Hall, .
- [51] H. Fu, M.K. Ng, M. Nikolova, and J.L. Barlow, "Efficient Minimization Methods of Mixed l2-l1 and l1-l1 Norms for Image Restoration," *SIAM Journal on Scientific Computing*, vol. 27, 2006, p. 1881.

- [52] C. Jaynes, A. Kale, N. Sanders, and E. Grossmann, "The terrascope dataset: A scripted multi-camera indoor video surveillance dataset with ground-truth," *Proceedings of the IEEE Workshop on VS PETS*, Citeseer, 2005, pp. 309-316.

Abstract

People Behavior analysis has lots of applications in various fields including customer behavior analysis in shopping centers and protection of facilities such as banks and airports. Tracking people is the first step in such systems. It is not possible for a single camera to observe a wide environment due to the limitation of sensors. Therefore a network of cameras is needed. In addition of Single camera tracking problems, Multi camera systems especially non-overlapping views encounter with other challenges. For example appearance of people is changeable due to situations such as different illumination, position of people, view angle and other parameters of the camera. In addition presence of people in camera views doesn't follow a certain rule and they can widely separate in time and space especially when cameras have no common area in their views, which make the problem more challenging.

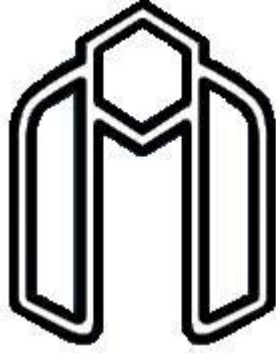
In this thesis, a new method is proposed to track people across disjoint camera views for an indoor environment. In this method, tracking people is performed in two stages. In the first stage people are tracked in each camera independently. Occurred events in each camera are saved in a central database with a standard data structure. In the second step, each person is tracked in whole environment, based on the information saved in database.

Background modeling and frame difference is used in order to detect people in a single camera. After applying pre-processes to improve the quality of detected objects and eliminating disturbing thing such as shadows, trajectory of each person is determined by Kalman Filter and MHT. Proposed method can track people in different illumination conditions and in presence of occlusions.

In order to establish correspondence between observations, which are related to a person, across cameras, Extracted features are compared with those **that** are saved in database. In this thesis a new feature is proposed that is more robust in different illumination conditions as well as being more distinguishable. This feature utilizes the main diagonal elements of Co-Occurrence matrix in YCbCr color space.

Proposed method is tested on samples from 5 cameras with different views of a building. Results showed that this feature is more robust than features like color.

Keywords: people tracking, Co-Occurrence matrix, Appearance model, Color space



Shahrood University of Technology

Faculty of Computer and IT Engineering

Tracking People Across Disjoint Camera Views for Indoor Environment

Rahman Yousefzadeh

Supervisor:

Dr. Hamid Hassanpour

Associate supervisor:

Dr. Ali.A Pouyan

January 2011