

دانشگاه صنعتی شاهرود

دانشکده: برق و رباتیک

گروه: الکترونیک

پایان نامه دوره‌ی کارشناسی ارشد مهندسی برق - الکترونیک

آشکارسازی و ردیابی تصویری وسایل نقلیه متحرک بر اساس

روش‌های پیش‌بینی

سالحه زهرودی

استاد راهنما:

دکتر امیدرضا معروضی

استاد مشاور:

دکتر علیرضا احمدی فرد

بهمن ۱۳۹۰



مدیریت تحصیلات تکمیلی
فرم شماره (۶)

شماره : ۰۹۳
تاریخ : ۹۰/۱۱/۱۶
ویرایش : -----

بسمه تعالی

فرم صورتجلسه دفاع پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) جلسه دفاع از پایان نامه کارشناسی ارشد خاتم / آقای :
صالحه زهرودی رشته : بسرق گرایش : الکترونیک
تحت عنوان : آشکارسازی و ردیابی تصویری وسایل نقلیه متحرک بر اساس روش های پیش بینی
که در تاریخ ۹۰/۱۱/۱۶ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح زیر است :

قبول (با درجه : عالی) امتیاز (۱۹٫۱۶۸) دفاع مجدد مردود

۱- عالی (۲۰ - ۱۹) ۲- بسیار خوب (۱۸/۹۹ - ۱۸)

۳- خوب (۱۷/۹۹ - ۱۶) ۴- قابل قبول (۱۵/۹۹ - ۱۴)

۵- نمره کمتر از ۱۴ غیر قابل قبول

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنما	ایدراف سرور	استادیار	
۲- استاد مشاور	عبدرضا احمدزاد	استادیار	
۳- نماینده شورای تحصیلات تکمیلی	سید علی حسینی	استادیار	
۴- استاد ممتحن	هادی کریمی	استادیار	
۵- استاد ممتحن	سعید حسینی	استادیار	

رئیس دانشکده :

تقدیم به:

پدر، مادر و همسر عزیزم که هستی و همه‌ی وجودم از آنهاست.

تشکر و قدردانی

در ابتدا از استاد راهنمای خود، جناب آقای دکتر امیدرضا معروضی به خاطر راهنمایی و حمایتی که از من در طی انجام این تحقیق کرده‌اند، نهایت تشکر را می‌نمایم. همچنین از استاد مشاور خود جناب آقای دکتر احمدی فرد تشکر می‌نمایم. از جناب آقای دکتر مروی و آقای دکتر سلیمانی که افتخار شاگردی کلاس درس آنها نصیب اینجانب گردید، کمال تشکر و قدردانی را دارم.

و نیز از تمامی دوستانم در دانشگاه صنعتی شاهرود، تشکر می‌نمایم، کسانی که بدون حمایت و حضور آنها، انجام این تحقیق، میسر نبود.

از خانواده عزیزم که در فراز و نشیب زندگی و تحصیل تنها تکیه گاه و پشتیبان محکم من بودند و همواره راهنمای دلسوزم و آرامش بخش لحظات زندگیم هستند، صمیمانه سپاسگزاری می‌نمایم و از خداوند متعال سلامتی آنها را خواستارم.

صالحه زهرودی

بهمن ۱۳۹۰

دانشجو تأیید می نماید که مطالب مندرج در این پایان نامه (رساله) نتیجه تحقیقات خودش می باشد و در صورت استفاده از نتایج دیگران مرجع آن را ذکر نموده است.

کلیه حقوق مادی مترتب از نتایج مطالعات ، آزمایشات و نو آوری ناشی از تحقیق موضوع این پایان نامه (رساله) متعلق به دانشگاه صنعتی شاهرود می باشد .

ماه و سال: بهمن ۹۰

یکی از مهمترین حوزه‌های پژوهشی در سیستم‌های حمل و نقل هوشمند، توسعه سیستم‌هایی است که جریان ترافیک در تقاطع‌ها را بصورت اتوماتیک مشاهده و کنترل می‌کنند. این سیستم‌ها علاوه بر تحلیل جریان حرکتی، می‌بایست تمامی جزئیات حرکتی هر وسیله نقلیه مانند موقعیت و سرعت را ثبت نمایند.

در این پایان‌نامه روشی جدید برای به حداقل رساندن خطاهای ناشی از فرآیند شناسایی و ردیابی خودرو در صحنه‌های شلوغ همانند تقاطع‌ها و جاده‌های درون شهری که همپوشانی بین خودروها و پیچیدگی حرکات بیشتر از بزرگراه‌ها و جاده‌های برون شهری می‌باشد، ارائه شده است. در روش پیشنهادی بمنظور تشخیص وسایل نقلیه متحرک از روش مدل ترکیبی گوسی جهت استخراج اشیاء متحرک موجود در تصویر ورودی استفاده می‌شود و در ادامه برای حذف غیرخودرو از خودروهای شناسایی شده و همچنین تقویت خروجی مرحله تشخیص از روش ماشین بردار پشتیبان، که به عنوان قوی‌ترین الگوریتم طبقه‌بندی مطرح می‌باشد استفاده خواهیم نمود. برای ورودی ماشین بردار پشتیبان نیاز به بردار ویژگی قدرتمندی برای تمایز خودرو از غیرخودرو داریم که نسبت به جهت و اندازه خودرو مستقل باشد. به همین منظور از توصیف‌گر هیستوگرام گرادیان‌های جهت‌دار (HOG)¹ استفاده می‌نماییم.

برای رفع مشکل همپوشانی خودروها، سیستمی جهت تشخیص همپوشانی پیشنهاد شده است که با استفاده از شکل ظاهری و مشخصات مورفولوژیکی خودروها، مکان همپوشانی را تشخیص داده و بمنظور جداسازی خودروهای داخل ناحیه همپوشانی، از یک مدل SVM که با استفاده از تصاویری با

¹ - Histogram of Oriented Gradients

اندازه‌های گوناگون آموزش دیده‌است، استفاده می‌نماید. سپس خروجی این مرحله به واحد ردیابی که بر پایه فیلتر کالمن و الگوریتم ردیابی فرضیات چندگانه^۱ استوار است وارد خواهد شد.

سیستم پیشنهادی ما می‌تواند، مشکلاتی از قبیل همپوشانی جزئی و تشخیص خودروها در صحنه‌های پیچیده را برطرف نماید. پیاده سازی‌ها با استفاده از پایگاه داده‌های مختلف در شرایط گوناگون جوی از تقاطع‌ها، مورد بررسی قرار گرفته است. خروجی نهایی سیستم فوق عملکرد مناسبی را با میانگین درصد تشخیص ۷۶/۹٪ ارائه داده است.

کلمات کلیدی: ردیابی و شناسایی خودرو - کنترل هوشمند ترافیک - مدل ترکیبی گوسی -

ماشین بردار پشتیبان - هیستوگرام گرادیان‌های جهت‌دار - فیلتر کالمن - ردیابی فرضیات چندگانه

^۱ Multiple Hypothesis Tracking (MHT)

لیست مقالات مستخرج از پایان نامه

Journal Paper: “Multiple Vehicle Detection and Tracking in Intersection and Crowded Scenes” S. Zohroudi, O.R. Marouzi, International Journal of Image Processing (IJIP), Volume 6, Issue 1.

Conference Paper: “Robust vehicle tracking in intersection and urban roads using histogram of oriented gradients”. S. Zohroudi, O.R. Marouzi, Proceedings of the IEEE, The Fourth International Conferences on Advances in Multimedia (MMEDIA 2012) Chamonix / Mont Blanc, France.

Conference Paper: “A new approach to multiple vehicle detection in intersection”. S. Zohroudi, O.R. Marouzi, Proceedings of the IEEE, 2012 International Conference on Electronics Computer Technology (ICECT 2012). Kanyakumari, India.

فهرست مطالب

صفحه	عنوان
۲	۱-مقدمه
۲	۱-۱- اهداف و انگیزش
۴	۲-۱-کاربردها
۵	۳-۱-تعاریف و مفاهیم پایه در ردیابی وسایل نقلیه
۶	۱-۳-۱-حسگرهای ترافیکی
۷	۲-۳-۱-روش‌های استخراج پارامترهای ترافیکی
۹	۴-۱-ساختار پایان‌نامه
۱۱	۲-روش‌های تشخیص وسایل نقلیه متحرک
۱۱	۱-۲-مقدمه
۱۲	۲-۲-بیان و نمایش اشیاء متحرک
۱۵	۳-۲-روش‌های تشخیص اشیاء متحرک
۱۷	۲-۳-۲-روش تفاضل فریم‌های متوالی
۱۹	۲-۳-۲-روش تفریق پس‌زمینه
۲۱	۲-۳-۲-تخمین پس‌زمینه با استفاده از مدل ترکیبی گوسی
۲۴	۲-۳-۴-روش جریان نوری
۲۶	۲-۴-۱-جداسازی و مدل‌سازی وسایل نقلیه متحرک
۲۷	۲-۴-۲-۱-جداسازی مبتنی بر مدل سه بعدی
۲۹	۲-۴-۲-۲-جداسازی مبتنی بر کانتورفعال
۳۰	۲-۴-۲-۳-جداسازی مبتنی بر ویژگی
۳۴	۲-۴-۲-۴-تطبیق و گروه بندی ویژگی‌ها
۳۵	۲-۴-۲-۵-جداسازی مبتنی بر ناحیه
۳۶	۲-۵-نتیجه گیری فصل

۳۷	۳-۱- روش‌های ردیابی وسایل نقلیه متحرک
۳۷	۳-۱-۱- مقدمه
۳۷	۳-۲- مروری بر روش‌های گذشته
۴۱	۳-۳- روش‌های ردیابی اجسام متحرک
۴۱	۳-۳-۱- ردیابی بر مبنای روش‌های قطعی
۴۳	۳-۳-۲- ردیابی بر اساس روش‌های پیش‌بینی
۴۴	۳-۴- فیلترینگ تصادفی
۴۶	۳-۴-۱- فیلترینگ بازگشتی بیزین غیر خطی
۴۹	۳-۵- کالمن
۵۴	۳-۶- ردیابی اهداف چندگانه
۵۶	۳-۶-۱- فرمول بندی بیزین عمومی برای ردیابی اهداف چندگانه
۵۸	۳-۶-۲- ردیابی فرضیات چندگانه (MHT)
۵۸	۳-۶-۳- الگوریتم JPDAF
۶۱	۳-۶-۴- فیلتر ذره‌ای
۶۴	۳-۷- نتیجه‌گیری فصل
۶۵	۴-۱- الگوریتم‌های پیشنهادی برای تشخیص و ردیابی
۶۵	۴-۱-۱- مقدمه
۶۸	۴-۲- تشخیص اولیه وسایل نقلیه متحرک
۶۸	۴-۲-۱- مدل ترکیبی گوسی و حذف نویز
۷۰	۴-۲-۲- حذف سایه از تصاویر مرحله تشخیص
۷۳	۴-۳- تقویت تشخیص وسایل نقلیه بر پایه روش‌های یادگیری
۷۵	۴-۳-۱- کلاسه‌بند SVM با استفاده از توصیف‌گر HOG
۸۴	۴-۳-۲- تشخیص همپوشانی
۸۹	۴-۴- نتایج آزمایشی
۱۰۷	۴-۵- نتیجه‌گیری فصل
۱۰۸	۵- نتیجه‌گیری و پیشنهادات

۱۱۰ ۲-۵- پیشنهادات برای کارهای آینده

۱۱۲ ۶-مراجع

شکل ۱-۲: روش‌های نمایش شیء (الف) یک نقطه (ب) چند نقطه (پ) ناحیه‌ی مستطیلی (ت) ناحیه‌ی بیضوی (ث) نواحی چندگانه که به یکدیگر وصل هستند (ج) اسکلت شیء (چ) ح) کانتورشیء (خ) سایه نما [۴]	۱۴
شکل ۲-۲: تشخیص خودرو با استفاده از روش تفاضل فریم‌های متوالی	۱۹
شکل ۳-۲: (الف) تصویر فریم تست (ب) تصویر حاصله از خروجی BS-GMM	۲۴
شکل ۴-۲: تصویر دو فریم متوالی، به همراه بردارهای جریان نوری و در نهایت تصویر متحرک جداسازی شده پایین سمت چپ نشان داده شده است. بردارهای حرکتی مربوط به جریان نوری نیز در تصویر سمت راست پایین نمایش داده شده است.	۲۶
شکل ۵-۲: یک نمونه مدل خودرو [۳۱]	۲۷
شکل ۶-۲: مراحل مختلف ساخت مدل خودرو [۳۱]	۲۸
شکل ۷-۲: نحوه همگرایی روش مرزفعال [۷]	۳۰
شکل ۸-۲: تاثیر عملگر لبه: ستون ۱، تصویر با سایزهای مختلف با استفاده از یک فیلتر پایین‌گذر بدست آمده است. در ستون ۲ و ۳ لبه‌های عمودی و افقی دیده می‌شود. در ستون ۴، پروفایل‌های عمودی و افقی تصاویر نشان داده شده است. ۳۲	۳۳
شکل ۹-۲: چهار گوشه برای تشخیص خودرو	۳۳
شکل ۱۰-۲: نمونه‌هایی از گروه‌بندی ویژگی‌ها [۳۷]	۳۵
شکل ۱-۳: عوامل محدودکننده در ردیابی قطعی وسایل نقلیه (الف) مجاورت (ب) حداکثر سرعت (پ) تغییرات جزئی سرعت (ت) حرکات مشترک (ث) ثبات شکل	۴۳
شکل ۲-۳: مراحل پیش‌بینی و بروز رسانی در تئوری بیزین	۴۸
شکل ۳-۳: مثال ارتباط اطلاعات. (الف) دو هدف (دایره‌ها) به همراه دو داده‌ی مشاهده شده (مثلث‌ها)، جهت ایجاد ارتباط. (ب) حالت ممکن اول در مسئله‌ی ارتباط اطلاعات. (ج) حالت ممکن دوم برای مسئله‌ی ارتباط اطلاعات. ... ۵۵	۵۵
شکل ۱-۴: شمای کلی الگوریتم پیشنهادی برای تشخیص و ردیابی وسایل نقلیه متحرک	۶۶
شکل ۲-۴: (الف) تصویر فریم‌های ورودی (ب) پس از مرحله حذف پس‌زمینه GMM. (ج) پس از حذف نویز و حذف نواحی کوچک و همچنین اعمال فیلترهای مورفولوژی	۶۹
شکل ۳-۴: (الف) تصویر فریم‌های ورودی در شرایط نویزی (ب) پس از مرحله حذف پس‌زمینه با استفاده از الگوریتم GMM. (ج) پس از حذف نویز و حذف نواحی کوچک و همچنین اعمال فیلترهای مورفولوژی	۷۰
شکل ۴-۴: (الف) تصاویر دو فریم ورودی (ب) پس از تشخیص اشیاء متحرک و تشخیص سایه	۷۲
شکل ۵-۴: (الف) تصاویر خروجی از مرحله تفریق پس‌زمینه GMM (ب تا ج) ردیابی در فریم‌های متوالی	۷۳
شکل ۶-۴: نمونه تصاویر آموزشی SVM (الف) مثبت (ب) منفی (پ) منفی تقریبی	۷۶
شکل ۷-۴: خلاصه‌ی روش HOG [۵۲]	۷۸

- شکل ۴-۸: اختصاص هیستوگرام به هر سلول برای دو نمونه تصویر ورودی ۸۰
- شکل ۴-۹: منحنی ROC مربوط به مدل آموزشی شماره ۱ با توجه به اندازه‌های مختلف فاکتور جهت ۸۱
- شکل ۴-۱۰: نمایش خروجی احتمالات مدل SVM برای تعیین میزان آستانه بین دو کلاس مثبت و منفی ۸۲
- شکل ۴-۱۱: (الف) تصاویر خروجی مرحله GMM (ب) تشخیص همپوشانی ۸۵
- شکل ۴-۱۲: تصاویر نمونه آموزشی با اندازه‌های متفاوت (a) مثبت (b) منفی (c) منفی تقریبی ۸۶
- شکل ۴-۱۳: منحنی ROC مربوط به مدل آموزشی شماره ۲ با توجه به اندازه‌های مختلف فاکتور جهت ۸۷
- شکل ۴-۱۴: نمونه تصاویری از ویدئوهای پایگاه داده استفاده شده در این پایان‌نامه ۹۰
- شکل ۴-۱۵: ردیابی بدون در نظر گرفتن سیستم مبتنی بر یادگیری (الف) خروجی GMM (ب تا ج) ردیابی در فریم‌های ۳۰۹ تا ۳۱۶ ۹۳
- شکل ۴-۱۶: (الف تا ت) ردیابی در چند فریم با استفاده از سیستم مبتنی بر یادگیری در روش پیشنهادی ۹۴
- شکل ۴-۱۷: ردیابی بدون در نظر گرفتن سیستم پیشنهادی (الف) خروجی GMM (ب تا ت) ردیابی در فریم‌های متوالی ۹۵
- شکل ۴-۱۸: ردیابی با اعمال سیستم پیشنهادی در دو فریم متوالی ۹۶
- شکل ۴-۱۹: (الف) خروجی GMM (ب تا ج) ردیابی بدون در نظر گرفتن سیستم با یادگیری و در نتیجه تشخیص و ردیابی ۲ خودرو بعنوان یک خودرو ۹۷
- شکل ۴-۲۰: ردیابی با اعمال سیستم پیشنهادی (الف تا ح) نتیجه تشخیص و ردیابی جداگانه خودروها در فریم‌های متوالی ۹۸
- شکل ۴-۲۱: ردیابی با در نظر نگرفتن سیستم پیشنهادی و عدم تشخیص همپوشانی در فریم‌های متوالی ۹۹
- شکل ۴-۲۲: تشخیص همپوشانی توسط سیستم پیشنهادی ۹۹
- شکل ۴-۲۳: (الف تا پ) ردیابی با اعمال سیستم پیشنهادی و تشخیص همپوشانی ۱۰۰
- شکل ۴-۲۴: (الف و ب) عدم تشخیص همپوشانی ۱۰۰
- شکل ۴-۲۵: تشخیص همپوشانی ۱۰۱
- شکل ۴-۲۶: (الف تا ت) نتیجه ردیابی در چند فریم متوالی با تشخیص همپوشانی ۱۰۱
- شکل ۴-۲۷: ردیابی نهایی توسط سیستم پیشنهادی در ویدئوهای مختلف ۱۰۲
- شکل ۴-۲۸: تشخیص نادرست همپوشانی ۱۰۳
- شکل ۴-۲۹: (الف و پ) خروجی GMM (ب و ت) خطای تشخیص و ردیابی در سیستم پیشنهادی ۱۰۵
- شکل ۴-۳۰: تشخیص نادرست همپوشانی با وجود اعمال سیستم پیشنهادی ۱۰۵

فهرست جداول

صفحه

عنوان

۸۳	جدول ۱-۴: ارزیابی مدل SVM شماره ۱
۸۸	جدول ۲-۴: ارزیابی مدل SVM شماره ۲
۱۰۵	جدول ۳-۴: نتیجه نهایی مربوط به ارزیابی کلی سیستم

فصل اول

مقدمه



۱-۱ - اهداف و انگیزش

با توجه به افزایش تعداد خودرو در جاده‌ها و همچنین ازدیاد تصادفات ناشی از تخلفات رانندگی، لزوم داشتن یک سیستم کنترل و نظارت احساس می‌شود. جوامع امروزی به شدت متکی به کار با ماشین‌های مختلف است و مسیر تکنولوژی آینده در جهت خودکار نمودن تمامی سیستم‌ها پیش می‌رود. نیاز هرچه بیشتر به سیستم ردیابی اشیاء با دقت، سرعت و قابلیت بیشتر و بالاتر بیش از پیش احساس می‌شود و در بسیاری از کاربردهای مراقبت ویدیویی و مانیتورینگ و بطور کلی کاربردهای بینایی ماشین، واحدی به‌منظور تشخیص و ردیابی، بعنوان یک پردازش مرتبه پایین‌تر نیاز می‌باشد. در حقیقت خروجی سیستم تشخیص و ردیابی که همانا اشیاء مورد ردیابی می‌باشند، بعنوان ورودی در پردازش‌های مرتبه بالاتر مانند تعبیر و تفسیر حرکت، شمارش تعداد اشیاء، تشخیص نوع رفتار و نظایر آن بکار می‌رود. امروزه کنترل هوشمند ترافیک در بسیاری از کشورها مورد توجه واقع شده است، برای مثال در امریکا با استفاده از تنها یک پنجم از کل بودجه سالیانه تخصیصی برای ساخت جاده‌ها، در بکارگیری کنترل هوشمند ترافیک، به همان کارایی مورد نظر رسیده‌اند.

کنترل یکپارچه و مناسب تجهیزات ترافیکی مانند علائم رانندگی و چراغ‌های راهنمایی تاثیر بسزایی در بهبود کنترل ترافیک دارد. بنابراین در برخی از کشورها به‌منظور نظارت و طراحی عملیات ترافیکی جاده‌ها و کاستن آمار تصادفات از مهندسی ترافیک بهره می‌برند. در این مهندسی، با مطالعه جریان ترافیک در جاده‌های مختلف، یک الگوی پایدار برای هر جاده به منظور روان کردن ترافیک آن، بدست می‌آورند. مهندسی ترافیک برای انجام طرح‌های خود نیاز به دانستن پارامترهای زیادی از یک جریان ترافیک در یک جاده دارد، که شامل؛ تشخیص تعداد و نوع خودروها، تعیین سرعت متوسط و لحظه‌ای خودروها، حجم ترافیک (تعداد خودرو که در واحد زمان از یک نقطه مشخص عبور می‌کند)، چگالی ترافیک (تعداد خودرو در واحد طول)، درصد اشغال خط (درصد زمانیکه یک نقطه خاص از

جاده با یک وسیله نقلیه اشغال می‌شود)، میزان تردد (تعداد خودروهایی که در واحد زمان از نقطه خاصی از یک مسیر عبور می‌کنند)، نرخ توزیع خودرو در تقاطع، می‌باشد

هدف از این تحقیق، تشخیص و ردیابی وسایل نقلیه متحرک در تقاطع می‌باشد. تعداد خودروها در تقاطع بالا بوده و وسایل نقلیه در جهات مختلفی حرکت می‌کنند، به همین دلیل کنترل هوشمند ترافیک در تقاطع از اهمیت ویژه‌ای برخوردار می‌باشد. اهداف متحرک در تقاطع‌ها شامل خودرو، موتور سیکلت و عابران پیاده می‌باشد، که در فرآیند تشخیص به تفکیک خودروها از دیگر اهداف متحرک می‌پردازیم. تصاویر ویدئویی استفاده شده در این تحقیق در شرایط آب و هوایی و مختلف و در ساعات گوناگون روز مورد بررسی قرار خواهد گرفت. همچنین فاصله و زاویه دید دوربین نسبت به خودروها متفاوت می‌باشد بنابراین روش‌های پیشنهادی این تحقیق تلاش در ارائه راهکاری کلی برای برطرف کردن مشکلات فوق دارد، بطوریکه ویژگی‌هایی که برای تشخیص وسایل نقلیه استفاده می‌شود نباید به موقعیت قرارگیری و زاویه دید دوربین وابسته باشد.

روش‌های پیشنهادی در این تحقیق برای تصاویر گرفته شده در طول روز مورد استفاده قرار می‌گیرند و تشخیص و ردیابی خودروها در شب از فرضیات این تحقیق نمی‌باشد. همچنین از آنجاییکه فرآیند تشخیص بکار گرفته شده از اطلاعات ناشی از چندین فریم متوالی بدست می‌آید، تشخیص خودروهای متوقف از فرضیات این تحقیق نمی‌باشد.

پیاده‌سازی الگوریتم‌های پیشنهادی توسط نرم افزار متلب و با استفاده از زیان برنامه نویسی C مورد بررسی قرار خواهد گرفت. خروجی بصورت شبه‌بلادرنگ است. برای پیاده‌سازی الگوریتم‌ها بصورت بلادرنگ نیاز به استفاده از پردازشگرهای DSP یا FPGA بوده که از اهداف این تحقیق نمی‌باشد.

۲-۱ - کاربردها

نظارت و کنترل ترافیک، شامل جمع‌آوری داده‌هایی است که خصوصیات خودروها و حرکت آنها را توصیف کند. معمولاً این داده‌ها در یکی از موارد زیر مورد استفاده قرار می‌گیرند: [۱]

- کنترل اجرای قانون، مانند کنترل حداکثر سرعت مجاز و یا تشخیص سبقت غیر مجاز و یا استفاده غیر مجاز از خط مخصوص اتوبوس
 - گرفتن اتوماتیک عوارض، که بدون متوقف کردن خودرو به طور اتوماتیک نوع آن را تشخیص داده و با توجه به نوع پلاک خودرو میزان عوارض مربوطه را در صورت حساب ماهیانه مشتریان ثبت می‌کند.
 - تشخیص تصادف و تراکم، که برای تسریع حضور نیروهای پلیس راهنمایی و رانندگی در محل استفاده می‌شود. علاوه بر آن با کمک علائم راهنمایی الکترونیکی، هشدارهای لازم جهت ایمنی رانندگان و جلوگیری از تصادفات ثانویه داده می‌شود.
 - افزایش ظرفیت راه: مسیریابی کردن اتوماتیک راه‌ها و چراغ قرمزها، به منظور کنترل ورودی بزرگراه‌ها با توجه به میزان ترافیک، در کاهش حجم ترافیک تاثیر بسزایی دارد.
 - حمایت از سیستم‌های حمل و نقل عمومی: ارائه اطلاعاتی شامل زمان حرکت و فاصله وسایل نقلیه عمومی تا ایستگاه، تعداد صندلی خالی و میزان کرایه، برای استفاده‌کنندگان از وسایل نقلیه بسیار سومند خواهد بود. لازم است این اطلاعات به صورت بلادرنگ جمع‌آوری شده و در اختیار متقاضیان قرار داده شود.
- از کاربردهای ردیابی در راهنمایی رانندگی می‌توان به تشخیص راننده خلافکار، استخراج شماره ماشین آن و جستجوی خودروهای مسروقه اشاره نمود.

۱-۳ - تعاریف و مفاهیم پایه در ردیابی وسایل نقلیه

در این بخش به معرفی برخی از مفاهیم پایه در سیستم تشخیص و ردیابی وسایل نقلیه می‌پردازیم که برای آشنایی بیشتر و سهولت فهم ادامه تحقیق مفید می‌باشد. هدف از ردیابی وسایل نقلیه، مشخص نمودن مکان آن‌ها در هر لحظه از زمان می‌باشد. ویژگی‌های مختلفی از وسایل نقلیه استخراج شده، و در فریم‌های بعدی برای یافتن محل جدید خودرو مورد استفاده قرار می‌گیرد. در صورتیکه خودروها به صورت پیوسته قابل مشاهده باشند، و ویژگی‌های آن‌ها در طول زمان تغییر نکند، مساله ردیابی، کار آسانی می‌باشد. اما در محیط‌های واقعی معمولاً، همپوشانی، تشخیص‌های ناقص و یا اشتباه در الگوریتم تشخیص خودرو، منجر به نتایج اشتباه می‌گردد.

در فرآیند ردیابی اشیاء متحرک، قبل از آن که ردیابی آغاز شود، ابتدا نیاز به تشخیص شیء مورد نظر و تفکیک یا جداسازی نواحی متحرک از پس‌زمینه داریم و سایر اطلاعات اضافی حذف می‌گردند. سپس ویژگی‌های منحصر بفرد شیء مورد نظر استخراج می‌گردد، ثبات ویژگی‌ها نسبت به شرایط محیطی، دقت بالا و هزینه محاسباتی پایین از جمله مواردی است که در این مرحله حائز اهمیت می‌باشد. سپس این ویژگی‌ها در هر فریم زمانی ردیابی می‌گردند.

با وجود اینکه تشخیص و ردیابی اشیاء متحرک برای چندین سال مورد مطالعه و تحقیق قرار گرفته است اما هنوز مباحثی باز برای تحقیق می‌باشد. اغلب مشکلات از متغیر بودن تصاویر ویدیویی به علت حرکت اشیاء ناشی می‌شود. هنگامی که یک شیء در زاویه دید دوربین حرکت می‌کند، تصاویر گرفته شده از شیء ممکن است، دچار تغییرات فاحشی شود. کاهش اطلاعات در اثر نگاهت از فضای سه‌بعدی به تصاویر دوبعدی هم مضاف بر علت می‌شود. نویز در تصاویر، تغییرات روشنایی صحنه، ابعاد متغیر اشیاء، همپوشانی با سایر اشیاء، پردازش بلادرنگ، جداسازی اشیاء از پیش‌زمینه و محیط از سایر مشکلاتی هستند که در فرآیند تشخیص و ردیابی اشیاء متحرک با آن‌ها مواجه هستیم.

۱-۳-۱ حسگرهای ترافیکی

نظارت جاده‌ها با استفاده از حسگرهای مختلفی صورت می‌گیرد، که در این بخش به معرفی مختصری از انواع حسگرها می‌پردازیم: [۲]

- حسگرهای تشخیص نیروی مکانیکی مثل آشکارسازهای هیدرولیکی یا پیزوالکتریکی
- آشکارهای آلتراسونیک و حسگرهای مادون قرمز که دقت بالایی ندارند و با ارسال امواج به سطح خیابان و مقایسه اختلاف زمان بین امواج منعکس شده از سطح وسایل نقلیه و امواج منعکس شده از سطح خیابان، وسایل نقلیه را تشخیص می‌دهند.
- رادار که معمولاً برای تعیین سرعت خودروها از روی تغییرات ایجاد شده در فرکانس امواج منتشره از سطح آن‌ها، استفاده می‌شود.
- حلقه‌های آشکارساز، که براساس تشخیص تغییرات انرژی در اثر وجود خودرو کار می‌کنند، مانند آشکارسازهای مغناطیسی و ماگنومتر
- حسگرهای تشخیص نور مثل دوربین‌های ویدئویی و چشم الکترونیکی

سنسورهای مغناطیسی، از دقت خوبی برای شمارش برخوردارند، اما نصب آن‌ها موجب تداخل در ترافیک شده و نگهداری از آن‌ها مشکل می‌باشد. سنسورهایی که در کف خیابان‌ها نصب می‌شوند، توسط برف‌روب‌ها و ... تخریب می‌گردند. عمل شمارش با استفاده از این تجهیزات زمانی که وسایل نقلیه گردش دارند با مشکل مواجه می‌گردند. برخی از تجهیزات قیمت بالایی دارند (لیزر) و برخی از دقت پایینی برخوردارند (ماکروویو). سنسورهای مادون قرمز در روز و شب عملکرد خوبی دارند و در آب و هوای همراه با مه، از سنسورهای با نور مرئی، عملکرد مناسب‌تری از خود نشان می‌دهند. سنسورهای آلتراسونیک، نسبت به شناسایی خودروهایی که با برف پوشیده شده‌اند، با مشکل مواجه است و همچنین به تغییرات دما و رطوبت حساس می‌باشد.

در میان انواع روش‌های معرفی شده، استفاده از بینایی ماشین برای نظارت ترافیک یکی از متداول‌ترین روش‌های کنترل ترافیک است و در بسیاری از موارد بهترین کارایی را دارد. و اطلاعات بدست آمده بسیار دقیق‌تر و شامل جزئیات بیشتری می‌باشد. یک نمونه ساده از یک سیستم کنترل ترافیک از طریق بینایی ماشین یک دوربین CCD معمولی است، که در مسیر حرکت خودروها نصب می‌شود و در ادامه با استفاده از تصاویر بدست آمده، الگوریتم‌های پردازشی مناسب برای تشخیص، ردیابی، دسته‌بندی و ... انجام می‌گیرد.

۲-۳-۱ روش‌های استخراج پارامترهای ترافیکی

در این بخش به بررسی روش‌های پردازشی، برای بدست آوردن پارامترهای مورد نیاز از یک صحنه ترافیکی می‌پردازیم. تنها ناحیه ثابت از تصویر که حاوی اطلاعات مهمی است، مورد بررسی قرار می‌گیرد. این کار باعث کاهش هزینه محاسباتی می‌گردد. [۳]

- روش حلقه مجازی: این روش در واقع شبیه‌سازی حلقه‌های مغناطیسی است که برای شمارش خودروها در زیر آسفالت خیابان نصب می‌شود. این الگوریتم بسیار ساده مشابه یک شمارنده عمل می‌کند و می‌تواند برای جداسازی خودروها از یکدیگر بکار گرفته شود. خطوط مشخصی در عرض جاده در نظر گرفته می‌شود، سپس با مقایسه مقادیر پیکسل‌های این خط، با مقادیر معادل آن در تصویر پس‌زمینه، وجود یک خودرو تشخیص داده می‌شود. ممکن است نیمی از خودرو و یا دو خودرو به طور همزمان از روی خط عبور کنند، و فرآیند شمارش را دچار مشکل نمایند. همچنین زمانی که شیشه‌جلو خودرو از روی این خط عبور می‌کند، هم‌رنگ زمینه بوده و ممکن است یک خودرو، دو بار شمرده شود.
- روش پنجره ثابت: در این روش یک پنجره مستطیلی ثابت روی مسیر عبور خودرو در نظر می‌گیریم، به طوری که کل باند جاده را پوشش دهد. نحوه پردازش پیکسل‌های موجود در

این پنجره، می‌تواند با استفاده از استخراج لبه و یا تفاضل‌گیری از پس‌زمینه باشد. در روش تفاضل‌گیری، پس از باینری کردن تصویر، از دو انتهای چپ و راست پنجره شروع کرده و به طرف مرکز پیش می‌رویم و در هر گام، روشنایی آن نقطه از تصویر باینری را با یک مقدار آستانه مقایسه می‌کنیم. از هر طرف اولین نقطه که تفاضل روشنایی آن با پس‌زمینه از مقدار آستانه بیشتر شد، مرز چپ یا راست را تشکیل می‌دهد.

روش دیگر استفاده از شناساگر لبه برای تشخیص وجود و یا عدم وجود خودرو در پنجره می‌باشد. مقدار آستانه‌ای که برای مقایسه تعداد پیکسل‌های لبه انتخاب می‌کنیم، بسیار حائز اهمیت می‌باشد. ممکن است با انتخاب یک مقدار آستانه بزرگ خودروهایی که هم‌رنگ جاده هستند را تشخیص ندهیم. برای حل این مشکل از تکنیک‌های آستانه‌گیری دوگانه و یا آستانه‌گیری تکراری، استفاده می‌کنند، که مقدار آستانه برای هر فریم با توجه به شرایط محیطی محاسبه می‌گردد. این روش عملکرد بهتری نسبت به روش حلقه‌های مجازی دارد، چراکه هر بار تعداد پیکسل‌های بیشتری را مورد بررسی قرار داده و مشکل قطعه‌قطعه شدن خودرو در این روش، کمتر اتفاق می‌افتد.

از مشکلات این روش، همپوشانی در ترافیک سنگین و عبور ناقص خودرو از پنجره می‌باشد، که سیستم را در تشخیص تعداد خودرو دچار مشکل می‌نماید.

با توجه به مشکلات مطرح شده، امروزه روش‌های ناحیه ثابت چندان کاربردی نداشته و معمولاً، نواحی متغیری در هر فریم مورد بررسی قرار خواهد گرفت. برای کاهش هزینه محاسباتی، تنها ناحیه‌ای را که احتمال وجود خودرو در آن بیشتر است، جستجو می‌نماییم. در فصل بعدی، مراحل مختلف سیستم‌های ردیابی مورد بررسی قرار خواهند گرفت.

۴-۱ - ساختار پایان نامه

این تحقیق شامل پنج فصل می‌باشد: مقدمه، کاربردها، مفاهیم پایه و تعریف مساله در فصل جاری مورد بررسی قرار گرفت.

در فصل دوم روشهای کلی تشخیص وسایل نقلیه بررسی می‌شوند. به علت اهمیت تشخیص اشیاء متحرک در سیستمهای مراقبتهای روشهای تشخیص وسایل نقلیه بر اساس حرکت با استفاده از اطلاعات چند فریم و بطور خاص روش حذف پس‌زمینه با استفاده از مدل ترکیبی گوسی مورد بررسی دقیقتر قرار می‌گیرند.

در فصل سوم پس از مرور مختصری بر روشهای موجود، در مورد نحوه‌ی استفاده از مدل‌های آماري و بطور خاص فیلتر کالمن و روش ردیابی فرضیات چندگانه برای پیش‌بینی موقعیت و ردیابی چندین وسیله نقلیه در تقاطع‌ها بحث خواهد شد.

در فصل چهارم سیستم پیشنهادی بصورت کامل تشریح خواهد شد. در این سیستم، جهت تشخیص وسایل نقلیه متحرک از روش مدل ترکیبی گوسی استفاده شده است و سپس عملیات حذف نویز و سایه و نواحی کوچک با استفاده از فیلترهای مختلف، انجام شده است. در ادامه برای حذف غیر خودرو از خودروهای شناسایی شده نیازمند بکارگیری یک سیستم یادگیری می‌باشیم که برای این منظور از ماشین بردار پشتیبان استفاده می‌کنیم، که به عنوان قوی‌ترین الگوریتم کلاسه بندی مطرح می‌باشد. برای ورودی ماشین بردار پشتیبان نیاز به بردار ویژگی بسیار قدرتمندی برای تمایز خودرو از غیر خودرو می‌باشیم که نسبت به جهت و اندازه خودرو مستقل باشد. به همین منظور از توصیف‌گر هیستوگرام گردیان جهت‌دار (HOG)^۱ استفاده می‌نماییم. خروجی این مرحله به واحد تشخیص و

^۱ - Histogram of Oriented Gradients

رفع همپوشانی وارد و سپس نتیجه حاصله به واحد ردیابی که بر پایه فیلتر کالمن و MHT استوار است وارد خواهد شد.

در فصل ۵ نیز یک جمع بندی کلی از آنچه در این پایان نامه آمده است خواهیم داشت و به نتیجه گیری از آنچه آورده شده است می پردازیم و در انتها راهکارهای آتی برای پژوهشهایی که در ادامه قابلیت انجام را دارند ارائه می دهیم.

فصل دوم

روش‌های تشخیص وسایل نقلیه متحرک

۲-۱ - مقدمه

در بسیاری از کاربردهای مربوط به پردازش تصویر، پیش از پرداختن به تحلیل‌های اصلی نیاز به تشخیص و جداسازی اشیاء پیش زمینه می‌باشد، از این رو تشخیص و آشکارسازی اشیاء متحرک، اهمیت زیادی در ردیابی اشیاء بخصوص در کاربردهای مراقبتی دارد، و حجم زیادی از تلاش‌های تحقیقاتی در این زمینه در دهه‌های گذشته به این مسئله معطوف گشته است.

در این فصل مروری بر روش‌های رایج در زمینه تشخیص اشیاء متحرک خواهیم داشت. مراحل مختلف این فرایند مورد بحث قرار گرفته و کارهای انجام شده در این زمینه را معرفی خواهیم کرد. در ادامه بحث به طور خاص روش‌های مطرح در تشخیص خودرو و کنترل ترافیک را مورد بررسی قرار می‌دهیم.

۲-۲ بیان و نمایش اشیاء متحرک

در یک الگوریتم ردیابی، در ابتدا باید شیء را به الگوریتم معرفی کنیم. نمایش‌های که استفاده از آنها در الگوریتم‌های ردیابی متداول هستند در ادامه معرفی می‌گردند. [۴]

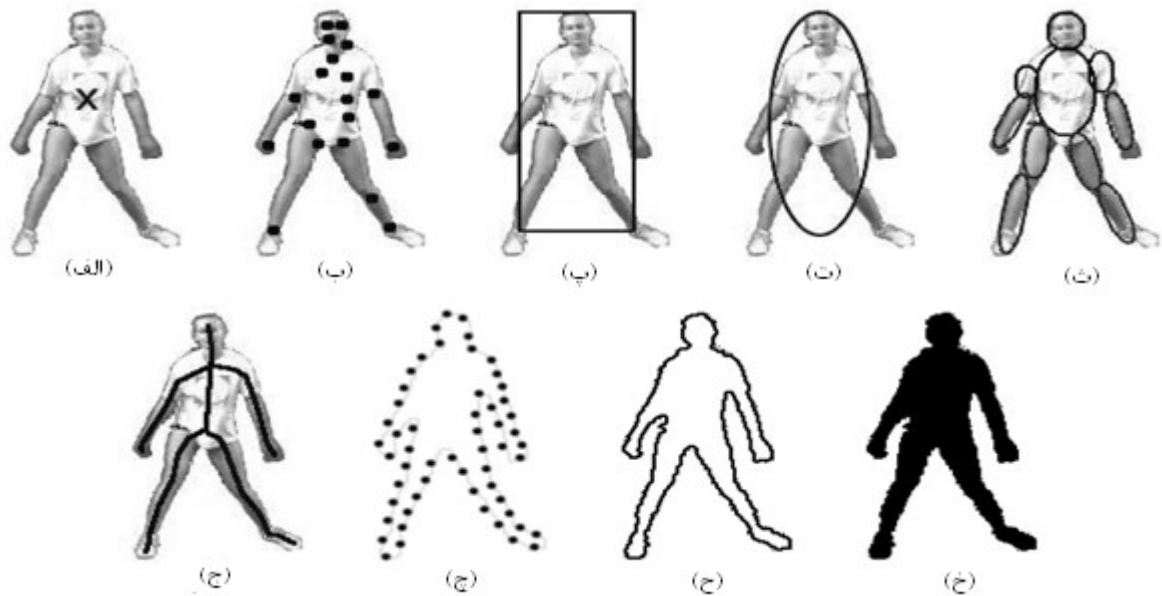
- نمایش توسط نقطه‌ها: شکل را می‌توان با یک نقطه (در مرکز ثقل جسم) (شکل ۱-۲ الف) یا مجموعه‌ای از نقاط (شکل ۱-۲ ب) نمایش داد. در حالت کلی نمایش با نقطه برای ردیابی اشیائی که ناحیه‌ی کوچکی از تصویر را اشغال می‌کنند مناسب است.
- نمایش با اشکال هندسی اولیه: شکل شیء با یکی از اشکال هندسی پایه مانند مستطیل (شکل ۱-۲ پ) یا بیضی (شکل ۱-۲ ت) قابل نمایش است. حرکت شیء برای چنین نمایش‌هایی معمولاً با تبدیل‌های هندسی پروجکتیو (هوموگرافی)، انتقال^۱، مدل می‌شود. اگرچه اشکال هندسی اولیه برای نمایش اشیاء غیرمفصلی^۲ مناسب‌تر هستند، لیکن برای ردیابی اشیاء مفصلی نیز به کار می‌روند.
- نمایش با کانتور و سایه نمای شیء^۳: نمایش کانتور، مرز شیء را تعریف می‌کند. (شکل ۱-۲ چ و ح) ناحیه‌ی درون مرز، سایه‌نمای شیء نامیده می‌شود. (شکل ۱-۲ خ) نمایش کانتور و سایه نما برای ردیابی اشیاء پیچیده‌ی مفصلی مناسب هستند. [۵]
- مدل‌های شکل بندبند^۴: اشیاء بندبند از قسمت‌هایی تشکیل شده‌اند که با اتصالاتی به هم متصل می‌شوند. برای مثال بدن انسان یک شکل بندبند است با تنه، پاها، دست‌ها و سر که با اتصالاتی در کنار هم قرار گرفته‌اند. (شکل ۱-۲ ث)
- مدل‌های اسکلتی^۱: این مدل نمایش را می‌توان برای مدل کردن اشیاء صلب و همچنین اشیاء بندبند مورد استفاده قرار داد. (شکل ۱-۲ ج)

¹ Affine

² Rigid

³ Object silhouette and counter

⁴ Articulated shape model



شکل ۱-۲ روش‌های نمایش شیء: الف) یک نقطه ب) چند نقطه پ) ناحیه‌ی مستطیلی ت) ناحیه‌ی بیضوی ث) نواحی چندگانه که به یکدیگر وصل هستند ج) اسکلت شیء چ) کانتور شیء خ) سایه نما [۴]

راه‌های مختلفی برای بیان ویژگی‌های ظاهر شیء وجود دارد. برای ردیابی، بیان شکل را می‌توان با بیان ظاهر ترکیب کرد. در ادامه بعضی از روش‌های متداول برای بیان ظاهر که در الگوریتم‌های ردیابی شیء به کار می‌رود آورده شده است:

چگالی احتمال ظاهر شیء^۲: چگالی احتمال ویژگی‌های ظاهر شیء (رنگ، بافت^۱) را می‌توان برای ناحیه‌ای از تصویر که با مدل‌های شکل (مثلاً ناحیه درونی یک بیضی یا یک کانتور) مشخص شده است، محاسبه کرد. روش‌های تخمین چگالی احتمال ویژگی‌های ظاهر شیء را می‌توان به دو دسته روش‌های پارامتریک مثل گوسین [۶] یا تلفیقی از گوسین‌ها [۷] یا غیرپارامتریک مثل پنجره‌های پارزن^۴ [۸] و هیستوگرام‌ها [۹] تقسیم نمود.

¹ Skeletal model

² Object appearance

³ Texture

⁴ Parzen windows

قالب‌ها^۱: الگوها توسط اشکال ساده هندسی یا سایه‌نما تشکیل می‌شوند [۱۰]. یک مزیت قالب این است که هم اطلاعات مکانی و هم اطلاعات ظاهر را با خود دارند. قالب‌ها ظاهر شیء را از یک نما در نظر می‌گیرند بنابراین آنها فقط برای مواردی که نمای مشاهده شده از شیء در طول فرایند ردیابی، تغییر چندانی ندارد، مناسب هستند.

مدل‌های نمایش فعال^۲: مدل‌های فعال ظاهر با مدل کردن هم‌زمان شکل و ظاهر شیء تولید می‌شوند [۱۱]. در این روش در حالت کلی شکل شیء توسط مجموعه‌ای از علامت‌های مشخصه^۳ تعریف می‌شود. علامت‌های مشخصه می‌توانند بر روی مرز شیء یا درون ناحیه شیء قرار گیرند. برای هر علامت مشخصه، یک بردار ظاهر^۴ که ویژگی‌های رنگ، بافت یا دامنه گرادیان را در بردار ذخیره می‌شود.

مدل‌های ظاهر چندنمایی^۵: این مدل‌ها نماهای مختلفی از یک شیء را در نظر می‌گیرند. یک راهکار برای نمایش نماهای مختلف شیء، تولید زیرفضایی از نماها می‌باشد. روش‌های زیرفضا بطور مثال تحلیل مولفه اصلی^۶ و تحلیل مولفه مستقل^۷ را هم برای بیان ظاهر و هم برای بیان شکل می‌توان به کاربرد [۱۲ و ۱۳].

¹ Templates

² Active appearance models

³ Landmarks

⁴ Appearance vector

⁵ Multiview appearance models

⁶ Principal Component Analysis (PCA)

⁷ Independent Component Analysis (ICA)

۲-۳ - روش‌های تشخیص اشیاء متحرک^۱

روشهای تشخیص اشیاء متحرک بر اساس تغییرات محتویات فریم‌های ویدیویی، اشیاء مورد نظر را از پس‌زمینه جدا می‌نماید. کارایی مراحل بعدی یک سیستم ردیابی، به آشکارسازی هدف بستگی دارد. تغییرات روشنایی، حرکات متناوب اجسام بر اثر وزش باد مانند برگ درختان، آشکارسازی را بامشکلاتی روبرو می‌سازد.

در تشخیص اشیاء می‌توان با استفاده از اطلاعات ناشی از یک فریم و یا چند فریم زمانی استفاده نمود. در روش‌هایی که تنها از یک فریم زمانی استفاده می‌کنند، نیاز به یک دانش اولیه در مورد شیء مورد نظر داریم. که خود به دو بخش تقسیم می‌شود:

- شناساگرهای نقطه‌ای^۲: که برای پیدا کردن نقاط مورد علاقه در تصاویر بکار می‌روند، مشخصات مورد علاقه نسبت به تغییر روشنایی و تغییر زاویه دید دوربین، مستقل هستند. مهمترین شناساگرهای نقطه‌ای شامل: عملگر Moravec's - شناساگر نقطه‌ای Harris [۱۴] و شناساگر SIFT^۳ [۱۵]
- شناساگرهای با سرپرست^۴:

تشخیص شیء متحرک را می‌توان با یادگیری نماهای مختلف شیء به کمک مجموعه‌ای از مثال‌ها انجام داد. یادگیری نماهای مختلف شیء نیاز به ذخیره‌سازی مجموعه کامل الگوها را از بین می‌برد. با استفاده از مجموعه‌ای از مثال‌های معدود یادگیری، این روش قادر است نماهای نادیده شیء را بیاموزد. در زمینه تشخیص شیء مثال‌های یادگیری از زوج‌هایی از ویژگی‌های شیء و کلاس شیء اختصاص داده شده به آن‌ها ترکیب شده‌اند

¹ Object detection

² Point detectors

³ Scale Invariant Feature Transform

⁴ Supervised Learning

که هر دوی این کمیت‌ها بصورت دستی تعریف می‌شوند [۴]. انتخاب ویژگی‌ها نقش مهمی در کارایی الگوریتم طبقه‌بندی ایفا می‌کند، بنابراین مهم است که از ویژگی‌هایی استفاده شود که یک کلاس را از کلاس دیگر تمیز دهند. هنگامیکه ویژگی‌های شیء انتخاب شدند، می‌توان یادگیری ظاهرهای مختلف از یک شیء را با استفاده از یک روش یادگیری با سرپرست انجام داد. شبکه‌های عصبی^۱ [۱۶]، بوستینگ تطبیقی^۲ [۱۷]، درخت‌های تصمیم^۳ [۱۸] و ماشین بردار پشتیبان^۴ [۱۹] از جمله روش‌های کارآمد در این زمینه می‌باشند. این روش‌های یادگیری، یک ابرصفحه^۵ که یک کلاس شیء را از کلاس‌های شیء دیگر جدا می‌کند، محاسبه می‌کنند.

روش‌های تشخیص اشیاء متحرک با استفاده از چند فریم، به سه دسته کلی شامل تفاضل فریم‌های متوالی، حذف پس‌زمینه و جریان‌نوری تقسیم می‌گردد. در دو روش اول فرض بر این است که، پس‌زمینه تصویر ثابت می‌باشد. در این دو روش صلب بودن یا غیر صلب بودن شیء اهمیتی ندارد، و در هر دو صورت شیء آشکار خواهد شد. اما روش جریان‌نوری در مواردی که پس‌زمینه ثابت نباشد، یا دارای حرکت کلی باشد نیز قابل استفاده است. ولی اشیاء در این روش باید صلب بوده و اجزاء مختلف شیء دارای مولفه‌های حرکتی ثابتی نسبت به مدل حرکتی پس‌زمینه باشند. علاوه بر این روش جریان‌نوری نسبت به دو روش قبل از بار محاسباتی بسیار بالاتری برخوردار است و اغلب از این روش بعلاوه بار محاسباتی سنگین، و حساسیت بالا نسبت به نویز، استفاده نمی‌شود. در ادامه این سه روش را با تفصیل بیشتری بیان خواهیم کرد.

¹ Neural networks

² Adaptive boosting

³ Decision trees

⁴ Support Vector Machine

⁵ Hypersurface

۱-۳-۲ روش تفاضل فریم‌های متوالی^۱

روش تفاضل فریم‌های متوالی نسبت به تغییرات نوری محیط تطبیق‌پذیر بوده و در برابر تغییرات زائدی که مابین دو فریم متوالی وجود دارد، و مطابق با تغییرات حرکتی شیء نمی‌باشد، از خود مقاومت نشان می‌دهد. [۲۰] این روش برای صحنه‌های پویا با تغییرات محیطی زیاد، بسیار مناسب عمل می‌کند، ولی در اکثر موارد در استخراج تمامی نقاط مربوط به شیء متحرک ضعیف عمل می‌کند، و معمولاً تنها نقاط مربوط به لبه‌های شیء را آشکار می‌سازد. مشکل عمده این روش، مشکل حفره‌ای شدن زمینه می‌باشد، بدین صورت که اگر سطح وسیعی از یک شیء متحرک، از نظر شدت روشنایی یکنواخت باشد، با وجود حرکت، وسط شیء ساکن به نظر می‌رسد و در تصویر تفاضل حفره‌ای ایجاد شده و موجب چندتکه‌ای شدن شیء می‌گردد. [۲۱ و ۲۲] همچنین تغییرات متوالی پیکسل‌های زمینه مانند حرکات برگ درختان، در این روش به اشتباه تشخیص داده می‌شود. و این روش برای صحنه‌های پیچیده مناسب نمی‌باشد. در این روش از تکنیک اعمال آستانه مناسب به تصاویر حاصل از تفاضل فریم‌های متوالی یک دنباله استفاده می‌شود. خروجی این روش فریم‌های دنباله را به دو ناحیه ثابت و متحرک (پس‌زمینه و پیش‌زمینه) تفکیک می‌کند. برای روشن شدن بحث رابطه (۱-۲) را در نظر بگیرید. که در آن $I_n(x)$ نشام‌دهنده شدت روشنایی فریم در نقطه x و در فریم n ام می‌باشد، آنگاه فریم تفاضل بصورت زیر تعریف می‌شود.

$$F_D(x) = I_n(x) - I_{n-1}(x) \quad (1-2)$$

فریم تفاضل در حقیقت تفاضل نقطه به نقطه بین دو فریم است. با فرض اینکه شرایط نوری محیط تقریباً بین دو فریم متوالی ثابت می‌باشد، مقدار فریم تفاضل در یک نقطه بیانگر تغییر در آن نقطه می‌باشد. مقدار غیر صفر تفاضل به دلیل حرکت اشیاء موجود در تصویر، و نیز نویز مشاهده بین مقادیر یک نقطه در دو فریم متوالی ایجاد می‌گردد. به‌منظور تمایز بین تغییرات ایجاد شده بوسیله

¹ Temporal differencing

² Aperture

نویز و تغییرات بر اثر حرکت اشیاء، از روش جداسازی مبتنی بر آستانه استفاده می‌کنیم قدرمطلق مقدار هر نقطه در فریم تفاضل با یک آستانه مقایسه می‌شود چنانچه این مقدار از این آستانه بیشتر باشد، آن نقطه بعنوان نقطه متحرک یا پیش‌زمینه تلقی خواهد شد، و چنانچه این مقدار از آستانه کمتر باشد، این نقطه به عنوان پس‌زمینه تلقی می‌گردد. در رابطه (۲-۲) با استفاده از یک ماسک باینری این مفاهیم را نشان داده‌ایم:

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

$T_n(x)$ آستانه تصمیم‌گیری برای فریم n ام در نقطه \mathbf{x} می‌باشد. این آستانه از توزیع آماری

تغییرات مقدار روشنایی یک نقطه و میزان تشخیص اشتباه نقاط بدست می‌آید.

در عمل ممکن است که نقاط منفرد بعلت وجود نویز در خروجی پردازش ایجاد شود. این نقاط را می‌توان بوسیله پیش‌پردازش حذف نمود. بعنوان مثال با محاسبه پیوستگی ۴تایی یا ۸تایی نقاط باینری، پیش‌زمینه را به گروه‌های متصل (حباب) تفکیک می‌کنیم و اگر تعداد نقاط در یک حباب (مساحت حباب) از تعداد مشخصی کمتر باشد، این گروه بعنوان نویز تلقی شده و حذف می‌گردد. در روش دیگر، به منظور حذف این نقاط منفرد از روش تفاضل تجمعی که نیاز به حافظه بیشتری در تشخیص حرکت دارد، استفاده می‌شود. برای این منظور فرض کنید، که $I_n(x)$ فریم جاری و $I_{n-1}(\mathbf{x}), I_{n-2}(\mathbf{x}), \dots, I_{n-N}(\mathbf{x})$ فریم ماقبل آن باشد، آنگاه با اعمال تفاضل بین فریم جاری و هر کدام از فریم‌های قبلی، N فریم تفاضل بدست می‌آوریم. پس از اعمال ماسک باینری به N فریم تفاضل، اگر تعداد نقاط روشن برای مختصات یک نقطه در N فریم تفاضل متوالی از حدی بیشتر شود، احتمال اثر نویز کمتر خواهد شد و آن نقطه بعنوان نقطه پیش‌زمینه در فریم باینری حاصل لحاظ خواهد شد. در کل، مشکل اصلی این روش میزان حافظه موردنیاز برای پیاده‌سازی می‌باشد. برای مثال برای دو فریم متوالی، حاصل بدین گونه محاسبه خواهد شد:

Binary Mask_n(x) =

$$\begin{cases} 1 & \text{if } (|I_n(x) - I_{n-1}(x)| > T_n(x)) \text{ and } (|I_n(x) - I_{n-2}(x)| > T_n(x)) \\ 0 & \text{otherwise} \end{cases} \quad (3-2)$$

در شکل ۲-۲ خروجی این روش برای آستانه ثابت نشان داده شده است. همانطور که دیده می‌شود، تنها نقاط مربوط به لبه‌های شیء متحرک بعنوان پیش‌زمینه آشکار شده‌اند، و در صورتیکه سرعت شیء کمتر باشد، این لبه‌ها باریک‌تر خواهند شد و هنگامی که حرکت شیء متوقف شود، این روش قادر به تشخیص شیء نخواهد بود.



شکل ۲-۲: تشخیص خودرو با استفاده از روش تفاضل فریم‌های متوالی

۲-۳-۲ روش تفریق پس‌زمینه^۱

در این روش از تفریق پس‌زمینه برای تشخیص اشیاء استفاده می‌گردد. تفریق هر تصویر از تصویر مرجع را محاسبه کرده و سپس نتیجه با یک مقدار آستانه مقایسه می‌گردد. این روش برای کاربردهای بلادرنگ مناسب می‌باشد. در رابطه (۴-۲) B_i پس‌زمینه مرجع، I_i فریم جاری و T ، آستانه می‌باشد

$$|B_i - I_i| > T \quad (4-2)$$

پس‌زمینه تصویر همواره ثابت نمی‌باشد و لازم است تطبیق داده شود. برخی از مواردی که باعث

ایجاد تغییرات در پس‌زمینه در شرایط مختلف می‌شوند عبارتند از:

¹ Background Subtraction

- تغییرات روشنایی، در طول روز و یا تغییر ناگهانی روشنایی در داخل ساختمان با روشن کردن لامپ و...
- حرکت کردن دوربین و تغییرات فرکانس بالای پس‌زمینه مانند: حرکت شاخه‌های درختان، موج دریا
- تغییر در موقعیت ماشین‌های پارک شده در یک خیابان و...

برای حل مشکلات فوق، عمل بروزرسانی پس‌زمینه با دخالت دادن اطلاعات گذشته فریم‌ها و تعیین مقدار مناسب برای هر پیکسل از پس‌زمینه با توجه به چندین فریم قبلی انجام می‌شود. بدین ترتیب که به ازای هر پیکسل، میانگین مقادیر معادل آن پیکسل در چند فریم گذشته محاسبه شده و جایگزین می‌گردد. بهتر است نرخ بروزرسانی پس‌زمینه با توجه به شرایط محیطی مثلاً زمان تعیین گردد (به‌عنوان مثال در ساعات نزدیک غروب یا طلوع نرخ بروزرسانی بیشتر باشد). می‌توان با توجه به رابطه (۵-۲) بروزرسانی این بروزرسانی صورت گیرد:

$$B_{i+1} = \alpha I_i + (1 - \alpha) B_i \quad (5-2)$$

که در آن α ، نرخ یادگیری و معمولاً برابر ۰.۰۵ انتخاب می‌گردد. در این روش، ممکن است هنگامیکه اجسام متحرک از یک ناحیه عبور می‌کنند، در مقدار میانگین روشنایی آن ناحیه تاثیر گذارده و هر چه سرعت و تعداد اجسام بیشتر باشد، تاثیرات نامطلوب بیشتری را خواهیم داشت.

روش‌هایی از جمله میانگین گوسی مداوم^۱ [۲۳]، فیلتر زمانی میانه^۲، مقادیر ویژه زمینه^۳ [۲۴]، برای حل این مشکلات پیشنهاد شده‌اند. روش ترکیبی گوسی^۴، می‌تواند بر مشکلاتی نظیر حرکات

^۱ Running Gaussain Average

^۲ Temporal Medium Filter

^۳ Eigen Background

^۴ Mixture of Gaussian

کوچک متناب و حرکات برگ درختان، غلبه کند. در این روش مقدار هر پیکسل با مجموعه‌ای از چند توزیع گاوسی مدل می‌گردد. و این توزیع‌ها در هر فریم، بروزسانی می‌شوند.

۳-۳-۲ تخمین پس‌زمینه با استفاده از مدل ترکیبی گاوسی

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. [۲۵] این مدل مانند یک مدل مخفی مارکف^۱ یک حالت است که تابع چگالی احتمال آن دارای چند توزیع نرمال می‌باشد. احتمال تعلق بردار X به یک مدل ترکیبی گاوسی دارای M توزیع نرمال به شکل زیر بیان می‌شود:

$$P(X | GMM) = \sum_{i=1}^M c_i \cdot N(\mu_i, \Sigma_i) \quad (۶-۲)$$

که در آن c_i وزن مدل گاوسی i و μ_i به ترتیب بردار میانگین و ماتریس کواریانس توزیع نرمال هستند. کواریانس مدل ترکیبی گاوسی معمولاً به صورت قطری در نظر گرفته می‌شود. گرچه امکان استفاده از ماتریس کامل نیز وجود دارد. رابطه ی (۶-۲) را می‌توان با استفاده از تابع چگالی احتمال نرمال به صورت زیر بیان نمود:

$$P(X | GMM) = \sum_{i=1}^M c_i \cdot \frac{1}{(2\pi)^{\frac{d}{2}} \left| \Sigma_i \right|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(x_i - \mu_i)^T \Sigma_i^{-1} (x_i - \mu_i)\right) \quad (۷-۲)$$

که در آن d ، بعد فضای ورودی است. برای به دست آوردن پارامترهای مدل ترکیبی گاوسی شامل وزن مدل گاوسی و میانگین و کواریانس توزیع‌ها از الگوریتم ماکزیمم نمودن امید ریاضی^۲ (EM) استفاده شده است.

^۱ Hidden Markov Model (HMM)

^۲ Expectation Maximization

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 می‌شود:

$$p(x_t) = \frac{1}{n} \sum_{i=1}^n k(x_t - x_i) \quad (۸-۲)$$

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

$$p(x_t) = \sum_{i=1}^k (w_{i,t} * f(x_t | \mu_{i,t}, \sum_{i,t})) \quad (۹-۲)$$

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 می‌باشد:

$$f(x_t | \mu_k, \sum_k) = \frac{1}{(2\pi)^{\frac{n}{2}} \left| \sum_k \right|^{\frac{1}{2}}} e^{-\frac{1}{2} (x_t - \mu_k)^t \sum_k^{-1} (x_t - \mu_k)} \quad (۱۰-۲)$$

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

$$\sum_k = \delta_k I \rightarrow \sum_k = \begin{pmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \sigma_g^2 & 0 \\ 0 & 0 & \sigma_b^2 \end{pmatrix} \quad (۱۱-۲)$$

مزیت این روش این است که برای نگهداری چگالی شدت رنگ یک پیکسل نیاز به نگهداری همه‌ی پیکسل‌های مشاهده شده نبوده و با دو متغیر میانه و میانگین می‌توان تابع چگالی آن را تشکیل داد.

جهت بروزرسانی مدل ترکیبی گاوسی^۱ لازم است که δ_i, μ_i, W_i که پارامترهای i امین تابع گاوسی می‌باشند در هر مرحله بروز رسانی شود. در صورتی که مقدار پیکسل X_t با k توزیع‌های مربوط به آن مقایسه و مقدار پیکسل در معادله‌ی (۱۲-۲) صدق کرد، پارامترهای ذکر شده باید بروز رسانی گردند.

$$\frac{X_t - \mu_i}{\delta_i} > 2.5 \quad (12-2)$$

مقدار $2/5$ یک حد آستانه‌ی تجربی است که توسط نویسندگان مقاله پیشنهاد شده است. [۲۵] یکی از ویژگی‌های اساسی این روش این است که هر گاه مقدار جدیدی مانند نویز در پس زمینه ظاهر شود باعث خرابی و از بین رفتن زمینه نمی‌شود. اگر هیچ یک از K توزیع با مقدار پیکسل جدید (X_t) تطبیق نداشتند، مقدار جدید جایگزین کمترین احتمال توزیع می‌گردد.

در هر مرحله هر یک از وزن‌ها از وزن قبلی تأثیر می‌پذیرد. همچنین توزیعی که با مقدار پیکسل جدید تطبیق داده شده است به اندازه α برابر بیشتر از وزن‌های دیگر تحت تأثیر قرار می‌گیرد.

$$W_{k,t} = (1 - \alpha)W_{k,t-1} + \alpha m(w_k)x_t \quad (13-2)$$

در رابطه‌ی (۱۳-۲) α نرخ یادگیری و m در صورت تطبیق مقدار پیکسل با انحراف معیار، مقدار ۱ و در غیر این صورت مقدار صفر بر می‌گرداند:

$$m(w_k | x_t) = \begin{cases} 1 & \text{if } X_t \text{ matches distribution } k \\ 0 & \text{otherwise} \end{cases} \quad (14-2)$$

^۱ Mixture model updating

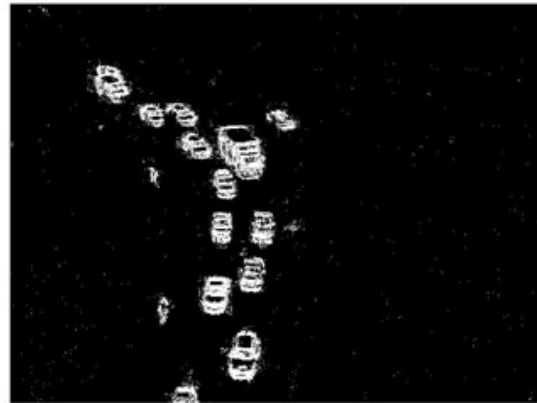
مقادیر میانگین و انحراف معیار نیز طبق فرمول‌های زیر به روز در می‌آیند:

$$\mu_{k,t} = (1 - f_k)\mu_{k,t-1} + f_k x_t \quad (15-2)$$

$$\delta_{k,t}^2 = (1 - f_k)\delta_{k,t-1}^2 + \eta_k (x_t - \mu_{k,t})^t (x_t, \mu_{k,t}) \quad (16-2)$$



(الف)



(ب)

شکل ۳-۲: (الف): تصویر فریم تست (ب) تصویر حاصله از خروجی *BS-GMM*

۴-۳-۲ روش جریان نوری^۱

روش جریان نوری، به فرم توزیعی از سرعت جابجایی روشنایی الگوها در یک تصویر تعریف می‌گردد و در آن جهت و سرعت اجسام متحرک به فرم بردار سرعت بدست می‌آید. از تکنیک‌های مشهور برای محاسبه جریان نوری می‌توان به [۲۹, ۲۸, ۲۷, ۲۶] اشاره نمود. همچنین ارزیابی مفصل روش‌های جریان نوری، در مرجع [۳۰] آمده است. این روش در موقعیت‌های تمیز دارای خروجی دقیقی می‌باشد. اما نسبت به تغییرات روشنایی بسیار حساس است و بار محاسباتی بالایی دارد. هرگاه جابه‌جایی برای یک پیکسل در دو لحظه زمانی داشته باشیم، بیانگر وجود یک جسم متحرک می‌باشد. در روش جریان نوری، برای هر پیکسل در همسایگی آن، تغییرات عمودی و افقی را در طول زمان با استفاده از تغییرات روشنایی در جهات عمودی و افقی بدست می‌آوریم. در این روش فرض می‌گردد که: (۱) - روشنایی هر پیکسل در دو فریم همسایه ثابت است (۲) - هر پیکسل حرکت

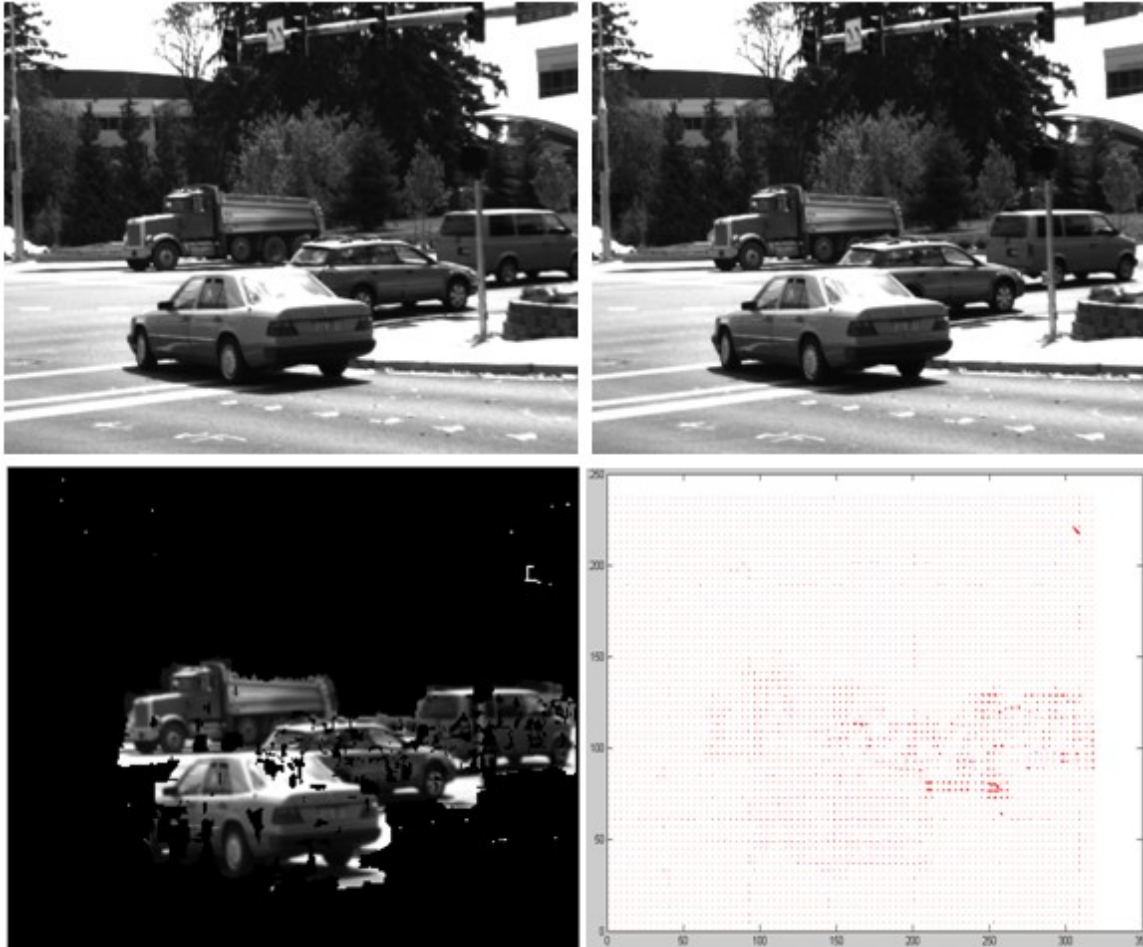
^۱ Optical Flow

کوچکی در هر فریم دارد به طوری که با بررسی همسایه‌های آن می‌توان بردار سرعت عمودی و افقی آن را بدست آورد. (۳) - تمام پیکسل‌های همسایه حرکتی مشابه با پیکسل مورد نظر دارند.

در ادامه بحث مثالی را از پیاده‌سازی روش Lucas-Kanade خواهیم داشت که برای جداسازی جسم متحرک از پس‌زمینه از یک مقدار آستانه مطابق رابطه (۲-۱۷) استفاده شده است. [۲۷]

$$Threshold = \sqrt{U^2 + V^2} \quad (۲-۱۷)$$

در این رابطه U, V بردارهای سرعت عمودی و افقی می‌باشند. در شکل ۲-۴ دو فریم متوالی از حرکت خودروها نشان داده شده است. در پیاده‌سازی از یک فیلتر پایین‌گذر برای حذف نویز استفاده شده است. که در هر مرحله‌ی فیلتر کردن اندازه تصویر چهار برابر کوچک می‌شود.



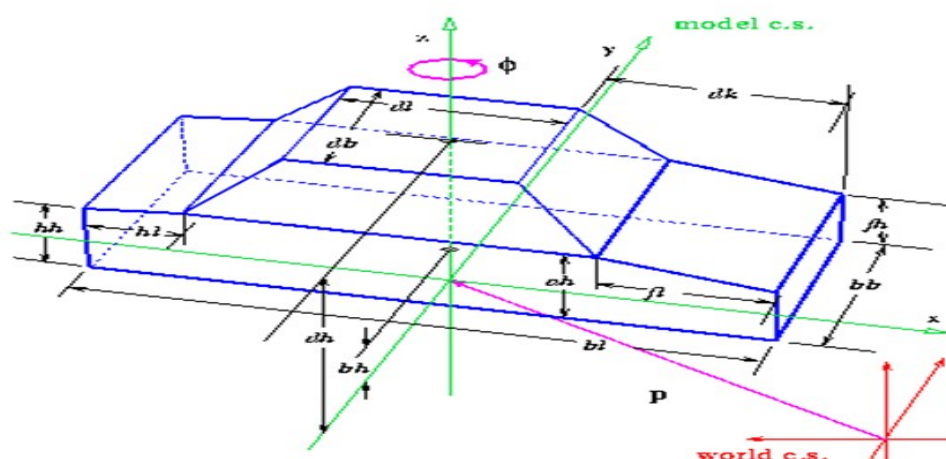
شکل ۲-۴: تصویر دو فریم متوالی، به همراه بردارهای جریان نوری و در نهایت تصویر متحرک جداسازی شده پایین سمت چپ نشان داده شده است. بردارهای حرکتی مربوط به جریان نوری نیز در تصویر سمت راست پایین نمایش داده شده است.

۲-۴ - جداسازی و مدل‌سازی وسایل نقلیه متحرک

در مرحله جداسازی و مدل‌سازی، نتیجه حاصل از مرحله تشخیص اشیاء متحرک بمنظور تفکیک خودروها از سایر اشیاء متحرک مورد پردازش قرار می‌گیرد. این تشخیص حتی می‌تواند تنها به صورت استخراج چند ویژگی بارز از هر خودرو باشد در ادامه با تعدادی از روش‌های متداول جداسازی جهت کاربردهای ترافیک آشنا می‌شویم.

۲-۴-۱ جداسازی مبتنی بر مدل سه بعدی^۱

در این روش، از یک یا چند مدل هندسی و دقیق از خودرو که قبلاً طراحی شده است استفاده می‌شود. مدل سه بعدی بر اساس این فرض طراحی شده است که خودرو یک جسم صلب و سخت بوده و دارای شکل ثابتی است و در جهت‌ها و فاصله‌های مختلف هم می‌توان، تصویر دوبعدی آن را، با دوران و تبدیل‌های مناسب از مدل اصلی، بدست آورد. [۳۱] با تصویر کردن انتقال‌یافته‌های مدل بر روی صفحه دو بعدی، بهترین تطبیق مدل با یک خودرو جستجو می‌شود. سپس بر اساس یک معیار شباهت میزان تطبیق مدل با خودروی موجود در تصویر اندازه‌گیری می‌شود، تا در نهایت مدلی که بیشترین تطبیق را داشته به عنوان مدل خودرو انتخاب و در فریم‌های بعدی ردیابی شود. روند کار به این صورت است که ابتدا خودرو از زمینه جداسازی شده، جهت آن مشخص می‌گردد. سپس خطوط بلند و صاف خودرو، استخراج شده و در ادامه پاره‌خطهای مدل خودرو برای قرار گرفتن در این جهت چرخانده شده و آنقدر جابجا می‌شوند تا دقیقاً بر پاره خط‌های تصویر منطبق شوند. در شکل ۲-۵ یک نمونه از این مدل‌ها را با ۱۲ پارامتر مختلف که شامل طول پاره خط‌های مدل و زاویه آنها است، مشاهده می‌کنید.



شکل ۲-۵: یک نمونه مدل خودرو [۳۱]

^۱ 3d Model Based

این روش در مقایسه با روش‌های مبتنی بر ویژگی، از نظر محاسباتی پیچیده‌تر و حجیم‌تر می‌باشد اما در رقابت مابین دقت و هزینه جانب دقت را می‌گیرد، به طوری که در اغلب پروژه‌هایی که این روش را انتخاب کردند، از چندین پروسسور قوی و به طور موازی استفاده می‌شود. همچنین این روش به مسأله روی هم افتادگی بسیار حساس است و تنها برای جاده‌های خلوت و با تراکم کم و ترافیک‌روان مناسب می‌باشد. در واقع روش‌های مبتنی بر مدل سه بعدی نتیجه تشخیص خودرو را به قیمت پرداخت هزینه محاسباتی زیاد با چنان دقتی ارائه می‌کنند که در اغلب کاربردها نیازی به آن همه جزئیات نیست. در شکل ۲-۶ مراحل مختلف این فرآیند مشاهده می‌گردد. ابتدا بر اساس میزان و جهت حرکت پیکسل‌ها، خودرو از زمینه جدا شده، جهت آن مشخص می‌گردد. سپس پاره‌خط‌های مدل خودرو برای قرار گرفتن در این جهت آنقدر جابجا می‌شوند تا دقیقاً بر پاره‌خط‌های تصویر منطبق شوند.



شکل ۲-۶: مراحل مختلف ساخت مدل خودرو [۳۱]

در کل روش‌های مبتنی بر مدل، به سطحی از فهم تصویر دست می‌یابند، در حالیکه در سایر روش‌ها تنها گروهی از پیکسل‌ها ردیابی می‌گردند.

روش‌های بر پایه مدل با استفاده از مدل لبه و تطبیق الگو مورد مطالعه قرار گرفته شده اند ولی در تصاویری با تعداد زیادی ماشین با اندازه‌های کوچک تطبیق و استفاده از مدل کار مشکلی است. در عمل تصاویر گرفته شده در اتوبان‌ها و جاده‌ها دارای درجه تفکیک بسیار پایینی هستند و استفاده از مدل‌های تطبیقی بسیار مشکل می‌باشد.

۲-۴-۲ جداسازی مبتنی بر کانتور فعال^۱

در روش‌های مبتنی بر کانتور (مرز) معمولاً از خودرو به عنوان یک شیء تغییرشکل‌پذیر، صحبت می‌شود. استفاده از عملگرهای ساده تشخیص لبه به تنهایی برای استخراج مرز چندان مناسب نیستند. چراکه این عملگرها، نمی‌توانند یک منحنی بسته به دور خودرو ایجاد کنند، و معمولاً لبه‌های گسسته تولید می‌کنند. علاوه بر آن لبه‌های درونی شیء را نیز مشخص می‌کنند، که چندان مناسب نمی‌باشد. یک راهکار اولیه استفاده از عملگرهای پیچیده‌تر تشخیص لبه مانند عملگر Canny است که در آنها لبه‌های مجاور و هم راستا بهم متصل می‌شوند. اما این عملگر نیز لبه‌های درونی را مشخص می‌سازد و همچنین از نظر زمانی، هزینه نسبتاً زیادی به سیستم اعمال می‌کند. [۷]

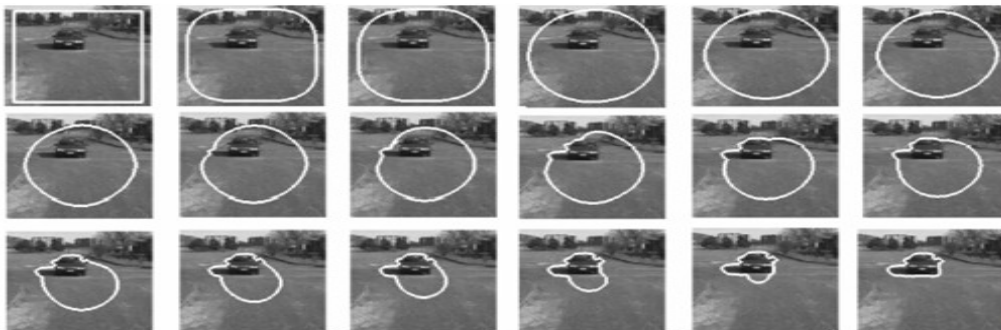
تبدیل هاف^۲، بر اساس رأی‌گیری از تعداد زیادی نقطه کار می‌کند، حساسیت زیادی به نویز ندارد، ولی در تصاویر پیچیده و با دقت پایین، که شکل خودرو در تصویر به صورت واضح و با خطوط مستقیم قابل تشخیص نباشد، مناسب نمی‌باشد.

برای تعیین مرز دور یک جسم از روشهای موسوم به مرز فعال استفاده می‌شود، که برای اولین بار در سال ۱۹۸۸ توسط Kass معرفی شد. یک مرز فعال، به کمک مجموعه‌هایی از نقاط محاسبه می‌گردد که یک تابع انرژی تعریف شده را، به حداقل می‌رسانند. [۳۲] برای یک مرز فعال دو تابع انرژی تعریف می‌گردد. انرژی داخلی، مسئولیت هموارسازی و متصل بودن و درکل پیکربندی شکل را بر عهده دارد و انرژی خارجی بر اساس نیروهایی است که از خارج از مرز برای جداکردن شیء از زمینه به آن تحمیل می‌شود. هدف از یک مرز فعال یافتن محلی برای نقاط مرزی است، که مجموع این انرژی‌ها را مینیمم کند. این روش چندان به کیفیت بالا در تصویر وابسته نیست، زیرا مرزها به صورت آزاد و غیر صلب در نظر گرفته می‌شوند و می‌توانند تغییر شکل هم بدهند. و از مشکلات این روش حساسیت بالایی است که به نقاط اولیه مرز دارد، و ممکن است در صورت تعیین نامناسب آن

¹ Snake

² Haugh Transform

نقاط، مرزفعال همگرا نشود و از مرز حقیقی شیء عبور کند، به ویژه اگر تعدادی از نقاط درون شیء در نظر گرفته شوند، احتمال یافتن یک مرز صحیح کمتر می‌گردد. همچنین انرژی خارجی برای نقاطی که از مرز فاصله دارند، نمی‌تواند نیروی زیادی اعمال کند و مرز مناسبی را تولید نمی‌کند. این روش تنها در ترافیکهای سبک به خوبی عمل می‌کند. مشکل دیگر روش مرز فعال نیز، هزینه محاسباتی بالای آن است و به علت ماهیت تکراری بودن آن برای کاربردهای بلادرنگی مانند نظارت بر ترافیک چندان مناسب نیست.



شکل ۲-۷: نحوه همگرایی روش مرزفعال [۷]

در شکل ۲-۷ مثالی از روش مرزفعال و مراحل همگرایی آن، نشان داده شده است

۲-۴-۳ جداسازی مبتنی بر ویژگی

در این روش بجای استخراج یک مدل کامل از خودرو که تمام تصویر آن را در بر گیرد، تنها چند ویژگی با معنا از آن را استخراج می‌کنیم، به این ترتیب ابتدا یک سری ویژگی‌های خاص از هر جسم استخراج شده، و سپس این ویژگی‌ها را بین تصاویر مختلف مقایسه می‌کنند. یک ویژگی به بخشی از تصویر می‌گویند که مشخصات موردنظر برای آن، نسبت به بیشتر بخش‌های مجاورش متفاوت باشد. [۳۳] ردیابی بر اساس ویژگی به سه دسته کلی تقسیم می‌شوند:

- ویژگی‌های عمومی^۱: ویژگی‌هایی که در این شاخه تعریف می‌شوند شامل مشخصه‌های عمومی مثل مرکز جسم، اطراف، مساحت، برخی مرتبه‌های توانی و رنگ‌ها می‌باشد.
- ویژگی‌های محلی^۲: ویژگی‌هایی که در این شاخه قرار می‌گیرند، از یک سری فاکتورهای شهودی مثل قطعه‌های خطی^۳ و منحنی وار^۴ و رئوس گوشه‌ها، تشکیل می‌شوند.
- ویژگی‌های خاص هندسی^۵: ویژگی‌هایی که در این حوزه قرار می‌گیرند یک سری ویژگی‌های خاص هندسی، می‌باشند. این روش حجم محاسباتی بالایی برای جستجو و مقایسه گراف‌ها بکار می‌گیرد و بلادرنگ نمی‌باشد

هدف اصلی در این مرحله جداکردن اجسام متحرک از یکدیگر و استخراج ویژگی‌های آن‌ها می‌باشد. در حالت کلی بهترین خصوصیت یک ویژگی تصویری منحصر به فرد بودن آن است، به گونه‌ای که اشیاء را بتوان به راحتی در فضای ویژگی از هم تشخیص داد. بسیاری از الگوریتم‌های ردیابی نیز ترکیبی از این ویژگی‌ها را به کار می‌برند. [۳۳] در زیر تعدادی از این ویژگی‌ها را مورد بررسی قرار می‌دهیم.

- رنگ: رنگ ظاهری یک شیء عمدتاً تحت تاثیر دو عامل فیزیکی توزیع طیف توان تابش کننده و ویژگی‌های بازتاب کنندگی سطح شیء قرار دارد. فضاهای رنگی متفاوتی از قبیل RGB, HSV, LUV, LAB داریم. در کل فضای رنگی نسبت به نویز حساس می‌باشد.
- لبه: لبه‌های جسم معمولاً تغییرات بارزی در شدت روشنایی تصویر ایجاد می‌کنند و روش‌های تشخیص لبه، برای نمایان ساختن این تغییرات بکار گرفته می‌شوند. عملگرهای معروف تشخیص لبه Sobel, Canny, Perwitt می‌باشند. در مورد یک وسیله نقلیه، با استفاده از زاویه‌های دید مختلف از یک اتومبیل، مخصوصاً زاویه دید از پشت و از روبرو می‌توان به

¹ Global Feature based Algorithm

² Local Feature based Algorithm

³ Line segment

⁴ Curve segment

⁵ Dependence Graph based

ساختارهای افقی و عمودی زیادی دست یافت. در شکل ۲-۸ تصاویری از یک خودرو با سایزهای مختلف دیده می‌شود. [۳۴]

- بافت: به تغییرات روشنایی ساختاریافته مانند صفحه شطرنج و یا تصویر پارچه طرح‌دار گفته می‌شود. در مقایسه با ویژگی رنگ، بافت نیاز به یک مرحله پردازش برای تولید توصیف‌گر دارد. توصیف‌گرهای مختلفی مانند: 'GLCM's^۱، موجک^۲، SIFT، آنتروپی و... وجود دارد. ویژگی لبه و بافت در مقایسه با ویژگی رنگ در مقابل تغییرات روشنایی حساسیت کمتری از خود نشان می‌دهند. [۳۵]



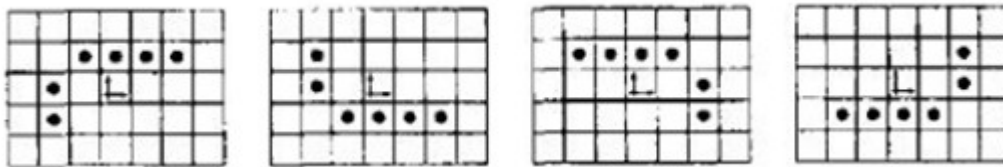
شکل ۲-۸: تاثیر عملگر لبه: ستون ۱، تصویر با سایزهای مختلف با استفاده از یک فیلتر پایین‌گذر بدست آمده است. در ستون ۲ و ۳ لبه‌های عمودی و افقی دیده می‌شود. در ستون ۴، پروفایل‌های عمودی و افقی تصاویر نشان داده شده است.

- تقارن (وسایل نقلیه): با استفاده از ویژگی تقارن به‌عنوان یک ویژگی مناسب در تشخیص اتومبیل استفاده نمود، زاویه دید از عقب یا جلوی اتومبیل در راستای افقی متقارن است.

^۱ Gray Level Co occurrence Matrics

^۲ Wavelet

- گوشه‌ها (وسایل نقلیه): اگر ساختار کلی هدف را شکل‌های مربعی در نظر بگیریم، می‌توانیم ۴ الگو برای گوشه‌ها طراحی کنیم که کاهش زمان جستجو برای تشخیص مکان خودرو، موثر است. در الگوریتم‌ها معیاری به نام درجه گوشه بودن^۱، وجود دارد، که اغلب از گرادیان مرتبه دوم استفاده می‌کنند و سپس مقایسه آن با یک مقدار آستانه، گوشه‌ها تشخیص داده می‌شوند. از میان عملگرهایی که برای تشخیص گوشه استفاده می‌شود، می‌توان Moravec, Harris, Smith را نام برد.



شکل ۲-۹: چهار گوشه برای تشخیص خودرو

- سایه: ناحیه زیر اتومبیل از آسفالت بسیار تیره تر از سایر نواحی روی آسفالت جاده می‌باشد، استفاده از این ویژگی برای کاهش فضای جستجو، مناسب می‌باشد.

در کل استفاده از ویژگی‌های فوق برای شناسایی وسایل نقلیه را می‌توان به صورت زیر خلاصه نمود: استفاده از ویژگی‌های تقارن، گوشه‌ها و بافت تنها در محیط‌های ساده و یکنواخت که در هم‌ریختگی‌های کوچک دارد، کارآمد است و در محیط‌های پیچیده، با خطای زیادی همراه است. در هوای مساعد، استفاده از اطلاعات سایه می‌تواند موثر باشد، هرچند که در وضعیت آب و هوای بد که پیکسل‌های تصویر تیره‌تر می‌شوند به شکست می‌انجامد. بکارگیری لبه‌های افقی و عمودی کارآمدترین رهیافت استفاده شده در روش‌های انتخاب ویژگی برای وسایل نقلیه می‌باشند. مشکل اصلی این رهیافت وابستگی آن به مقدار آستانه است که می‌تواند کارایی و قدرت سیستم را تحت الشعاع قرار دهد. ضمن اینکه روش‌های وابسته به لبه برای فواصل دور چندان مناسب نمی‌باشند.

¹ Cornerness

همچنین استفاده از ویژگی، در شب که در آن خودرو به عنوان یک جسم صلب دیده نمی‌شود و تنها چراغ‌ها و نقاط خاصی از آن واضح است، مناسب می‌باشد. مشکل همپوشانی در این روش کم‌رنگ‌تر شده است، چراکه اگر دو خودرو در فاصله بسیار نزدیک از هم قرار گرفته باشند، باز هم تنها در صورت پیدا بودن قسمت کوچکی از خودرو می‌توان چند ویژگی از آن بدست آورد.

۴-۴-۲ تطبیق و گروه بندی ویژگی‌ها

پس از یافتن ویژگی‌ها در هر فریم، باید آنها را با معادل آن‌ها در فریم‌های بعدی تطابق^۱ دهیم. یکی از روش‌های مطرح تطبیق‌دهنده بردار گرادیان^۲ می‌باشد، که در آن شدت روشنایی پیکسل گوشه و همچنین گرادیان‌های X, Y ویژگی با تمام ویژگی‌های کاندیدا در محدوده پیش‌بینی‌شده در فریم بعدی مقایسه می‌شود. و برای هر مشخصه یک بردار مشخصه، تعریف می‌گردد. سپس میزان شباهت بین یک ویژگی و ویژگی‌های کاندیدای آن به صورت محاسبه تفاوت بردار ویژگی در فریم t (که با W نشان داده شده) و فریم بعدی (با v نشان داده شده) با توجه به رابطه زیر محاسبه می‌شود:

$$m(v, w) = \frac{|v - w|}{\sqrt{|v| |w|}} \quad (18-2)$$

یکی دیگر از این روش‌ها، تطبیق‌دهنده ضریب ممان حاصل ضرب^۳ (PMCM) می‌باشد.

پس از تشخیص و تطبیق ویژگی نوبت به آن می‌رسد که ویژگی‌هایی را که به یک خودرو تعلق دارند با هم در یک گروه قرار دهیم و کل مجموعه را دنبال کنیم. برای گروه‌بندی ویژگی‌ها معمولاً از اطلاعات حرکتی، عملگرهای مشابهت و اصل چسبندگی اشیاء استفاده می‌شود. اصل چسبندگی بیان می‌کند که اجزای یک شیء در کنار هم و در فاصله ناچیزی از هم باقی می‌مانند. [۳۶]

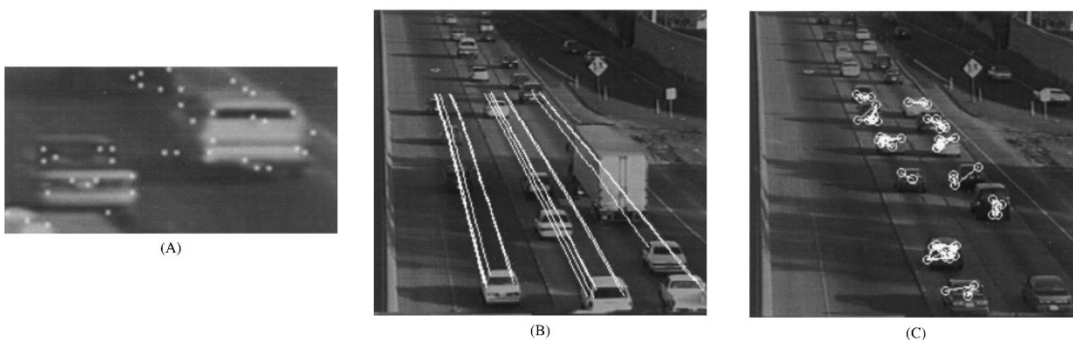
برای این کار معمولاً از یک گراف گروه‌بندی استفاده می‌شود. در این گراف گره‌ها، معادل با

¹ Correspondence

² Gradient Vector Matcher

³ Product Moment Coefficient Matcher

ویژگی‌های یافت‌شده و لبه‌ها به معنای هم‌گروه بودن است، یعنی دو گره مجاور به یک خودرو تعلق دارند. هرگاه یک ویژگی جدید تشخیص داده شود، به عنوان یک گره به گراف اضافه شده و به کلیه ویژگی‌های همسایه خود در یک شعاع معین متصل می‌شود. به عبارت دیگر فرض اولیه این است که کلیه ویژگی‌هایی که در یک شعاع همسایگی مشخص از هم قرار دارند به یک خودرو متعلق می‌باشند. در مرحله بعدی یک محدودیت به این گراف اضافه می‌شود تا خودروهای مجاور یکدیگر (که ویژگی‌های آن‌ها نیز در همسایگی هم هستند) از هم تفکیک شوند. به عنوان مثال می‌توان برای معیار ترکیب یا جداسازی ویژگی‌ها فاصله آن‌ها را در نظر گرفت، بدین معنی که اگر فاصله از یک میزان آستانه بیشتر بود، اتصال آن‌ها حذف گردد.



شکل ۲-۱۰: نمونه‌هایی از گروه‌بندی ویژگی‌ها [۳۷]

۲-۴-۵ جداسازی مبتنی بر ناحیه

در این روش بجای استخراج یک مدل دقیق، با تصویر خودرو به صورت یک ناحیه بهم پیوسته برخورد می‌شود. معمولاً پس از مرحله تشخیص خودرو، با اعمال فیلتر اندازه، قطعاتی که اندازه آن‌ها در حد اندازه یک خودرو نباشد، حذف می‌شود. [۳۸] سپس ویژگی‌ای مانند مرکز ثقل ناحیه و یا مستطیل از محیطی آن ناحیه را انتخاب می‌کنند. این شاخه بر اساس نواحی‌ای که در یک تصویر در جریان حرکت دچار تغییر می‌شوند، شیء را ردیابی می‌کند. همانگونه که از ظاهر این روش برمی‌آید، تمرکز الگوریتم بر اساس تغییراتی که به واسطه‌ی حرکت شیء در تصویر حاصل می‌گردد، استوار است. ویژگی‌هایی مانند رنگ و بافت از جسم استخراج می‌گردد. از مزایای این روش، فهم و پیاده‌سازی

آسان، معمولاً دقیق، و می‌تواند مشکل همپوشانی را اداره نماید. و از معایب این روش، حجم محاسباتی بالایی دارد و تا حدود زیادی به مرحله تشخیص وابسته است.

۲-۵ - نتیجه گیری فصل

در این فصل، زیر سیستم تشخیص و جداسازی وسایل نقلیه متحرک که اولین مرحله از یک سیستم ردیابی است مورد بررسی قرار گرفت. ابتدا برخی از روش‌های معمول از جمله تفاضل فریم‌ها، تفریق پس زمینه و شار نوری برای تشخیص وسایل نقلیه متحرک توضیح داده شده و مزایا و معایب هر یک ذکر گردید. همچنین مدل ترکیبی گاوسی به عنوان یکی از روش‌های موفق برای مدل‌سازی زمینه مورد بررسی قرار گرفته و نحوه‌ی آشکارسازی اهداف با آن توضیح داده شد، که بعلاوه کارایی بالا در سیستم‌های تشخیص وسایل نقلیه متحرک، از این روش در الگوریتم پیشنهادی استفاده خواهد شد. سپس روش‌های جداسازی وسایل نقلیه متحرک از سایر اشیاء متحرک شناسایی شده شرح داده شد.

فصل سوم

روش های ردیابی وسایل نقلیه محرک

۳ -۱- مقدمه

ردیابی یک هدف در یک ویدئو به معنای تطبیق اجسام شناسایی شده در فریم حال، با اجسامی که در فریم های قبلی شناسایی شده بودند، می باشد. بدست آوردن یک مسیر حرکتی در طول زمان برای این اجسام را ردیابی می گویند. همپوشانی، عدم شناسایی برخی از اجسام در یک فریم و اضافه یا کم شدن اجسام در فریم های متوالی، از مشکلات اصلی فرآیند ردیابی می باشد. در این فصل، پس از مرور مختصری بر روشهای موجود، در مورد نحوه ی استفاده از مدل های آماری برای پیش بینی موقعیت مربوط به یک وسیله نقلیه بحث خواهد شد.

۳ -۲- مروری بر روش های گذشته

- در سال ۱۹۹۹ مدلی برای زمینه با استفاده از یک مدل گاوسی مخلوط^۱ ارائه شد. [۲۵] مدل گاوسی مخلوط دارای انعطاف پذیری بالایی است و تغییرات ناگهانی در زمینه را به خوبی مدل

^۱ Gaussian Mixture Model

می‌کند. در ادامه با تفریق زمینه از تصویر هر فریم و آستانه‌ای کردن آن نواحی مربوط به اشیاء متحرک موجود در تصویر شناسایی شده است.

- در سال ۲۰۰۱ سامانه‌ی نظارتی دتر، جهت تشخیص رویدادها برای ارزیابی خطر و شناسایی ارائه شد. [۳۹] این سامانه به طرز قابل اطمینانی انسان‌ها و وسایل نقلیه را در شبانه روز مورد ردیابی قرار می‌دهد. در این سامانه از واحدهای بینایی ماشین و ارزیابی خطر استفاده شده است. در واحد بینایی ماشین شیء متحرک ردیابی می‌گردد. برای این منظور از چند دوربین در فضای سه‌بعدی استفاده شده است. توسط واحد ارزیابی خطر، الگوهای مظنون تشخیص و گزارش داده می‌شود. از این سامانه در سوپرمارکت‌ها و مکان‌های تجاری و امنیتی استفاده می‌گردد.

- در سال ۲۰۰۲ یک سامانه نظارت ویدئویی بلادرنگ پیاده سازی شده است که به تشخیص و ردیابی اشیاء متحرک در یک محیط باز و وسیع با استفاده از شبکه‌ای از دوربین‌ها می‌پردازد. [۴۰] سپس اشیاء متحرک را با استفاده از تفریق هر فریم از تصویر ثابت زمینه می‌یابد، و در ادامه با استفاده از فیلتر کالمن به تخمین حالت هر شیء می‌پردازد. در این سامانه با استفاده از چندین دوربین و در نتیجه داشتن چند زاویه دید مختلف از اجسام متحرک، مسئله همپوشانی برای هر کدام از اجسام را مورد بررسی قرار داده است.

در ادامه، مروری بر کارهای انجام شده در زمینه سیستم‌های کنترل ترافیک خواهیم داشت. [۲]

- سیستم TRIP: این پروژه که با همکاری دانشگاه منچستر و دانشگاه شفیلد در اوایل دهه ۸۰ با استفاده از تکنیک‌های بینایی ماشین پیاده‌سازی شد. در آن برای شمارش خودروها، از تفاضل تصاویر از تصویر جاده بدون خودرو به عنوان تصویر مرجع استفاده می‌کرد. مطابق گزارشات درصد خطا برای شمارش خودروها ۱٪ و برای محاسبه سرعت بین ۸٪ تا ۱۷٪ بوده است.

- سیستم TULIP¹: این سیستم در سال ۱۹۸۸ در انگلستان و با هدف شمارش تعداد خودروها تعریف و پیاده سازی شد. این سیستم به دو روش زیر اجرا می شد: روش اول باینری کردن تصاویر با یک سطح آستانه از پیش تعیین شده و جدا کردن نقاط روشن خودرو، و روش دوم به صورت اعمال چند سطح آستانه که شدت روشنایی سطح جاده و منطقه سایه را مشخص می کنند، بود. نتایج پیاده سازی این سیستم خطای متوسط ۳٪ تا ۱۳٪ برای شرایط مختلف آب و هوایی داشته است.
- سیستم CLIP4: این سیستم به صورت چند پردازنده ای و با پردازش موازی در یکی از دانشگاه های انگلستان پیاده سازی شده است. هدف از این سیستم ردگیری خودرو از زمینه با عملیات تفاضل گیری و ردگیری خودرو می باشد. توانایی این سیستم در حد ردگیری حداکثر ۵ خودرو در صحنه است.
- سیستم های SCATS و SCOOT: سیستم SCOOT در دهه ۷۰ در انگلیس ارائه شد و از جمله سیستم های کنترل ترافیک پرطرفداری است که بر اساس آمارها و نتایج لحظه ای ترافیک، در مورد زمان بندی چراغ های راهنمایی تصمیم گیری می کند. اثرات مثبت آن در مدت زمان سفر، تاخیر در سفر، آلودگی هوا و مصرف سوخت می باشد. سیستم SCATS به صورت لحظه ای عمل کرده و چراغ های راهنمایی را در کل سیستم در پاسخ به تغییرات در تقاضاهای ترافیک و ظرفیت سیستم تنظیم می کند. این سیستم جامع در استرالیا و ۵۰ نقطه دیگر جهان به کار گرفته شده است.
- تحقیقات در دانشگاه توکیو: در اوایل دهه ۸۰ پروژه ای در توکیو برای شمارش و اندازه گیری سرعت وسایل نقلیه انجام گرفت. در این پروژه از تعدادی نقاط نمونه در هر باند از جاده برای شناسایی خودرو استفاده شد. یعنی به جای پردازش کل تصاویر تنها بخش ثابت و کوچکی کنترل می شد. برای تعیین سرعت از دو نقطه نمونه در هر باند و

¹ Traffic Analysis Using Low Cost Image Processing

بررسی میزان تاخیر خودرو بین این دو نقطه کمک گرفته شد. نرخ خطای این پروژه برای شمارش خودرو ۵٪ و برای سرعت، حداکثر ۱۰٪ بوده است.

- تحقیقات در سوئد: موسسه صنعتی سلطنتی سوئد، پروژه‌ای برای شمارش و تعیین سرعت خودروها در چندین باند تعریف کرد. در این پروژه برای هر باند از یک پروفایل موازی با جاده، به پهنای سه پیکسل استفاده شد و مقادیر پیکسل‌های هر پروفایل با تصویر مرجع مقایسه شده است و از روی آن خودروهای مختلف جداسازی و دنبال می‌شدند. به این ترتیب تعداد و سرعت خودروها با مقایسه تعداد کمی پیکسل محاسبه می‌گردید.

- تحقیقات در کالیفرنیا: موسسه صنعتی کالیفرنیا پروژه‌ای برای تشخیص، شمارش و اندازه‌گیری سرعت خودروها به صورت بلادرنگ طراحی کرد که ارزیابی هر تصویر را بر روی یک خط افقی به طول ۵۰ پیکسل برای هر باند انجام می‌داد. ردیابی خودرو با بررسی ۳۰ پروفایل خطی که بر روی جاده در نظر گرفته شده بود، امکان‌پذیر می‌باشد. خطای این پروژه، برای شمارش ۴٪ و برای سرعت ۲٪ بوده است.

- سیستم IMPACTS: این سیستم که برای نظارت‌های کیفیتی ترافیک طراحی شده، در هر باند جاده، به تعدادی سلول متناسب با عرض جاده و طولی معادل، طول یک اتومبیل تقسیم می‌شود. برای هر سلول یکی از سه حالت خالی از خودرو، خودرو در حال عبور و خودرو متوقف تشخیص داده می‌شود. حالت هر سلول بر اساس تعداد لبه‌ها و همچنین تشخیص حرکت با تفاضل‌گیری فریم‌ها، مشخص می‌گردد. خطای سیستم در دسته بندی سلول‌ها، با توجه به شرایط آب و هوایی بین ۵٪ تا ۱۲٪ و خطای تشخیص تراکم (برای خودروهای متوقف) تنها ۳٪ بوده است.

- دانشگاه برکلی: این سیستم قادر به تشخیص و ردگیری همزمان چندین خودرو و تشخیص شکل آن‌ها برای دسته‌بندی انواع خودروهاست. در این پروژه هم از تفاضل فریم‌ها از فریم مرجع استفاده شده است. فریم مرجع با استفاده از فیلتر کالمن بروزرسانی

می‌شود. از خروجی این تفاضل برای مقدار دهی اولیه به منحنی مرزی^۱ برای هر خودرو استفاده شده‌است. برای پیاده‌سازی بلادرنگ از شبکه‌ای از پردازنده‌های دیجیتال استفاده شده است.

۳-۳ - روش‌های ردیابی اجسام متحرک

پس از اینکه اهداف 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 کاربرد های مختلف ارائه شده‌اند. روش‌های ردیابی وسایل نقلیه را می‌توان به دو دسته‌ی روش‌های قطعی و روش‌های آماری 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 (پیش‌بینی)، 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 می‌شود. حالات اشیاء در تصویر مانند ظاهر، موقعیت، سرعت و یا سایر مشخصات بر اساس حالات تشخیص داده 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.

۳-۳-۱ - ردیابی بر مبنای روش‌های قطعی

روش‌های قطعی برای به دست آوردن همبستگی بین نقطه‌های موجود در فریم N و $N+1$ از یک تابع هزینه که با استفاده از محدودیت‌های حرکت به دست می‌آید، استفاده می‌نمایند.

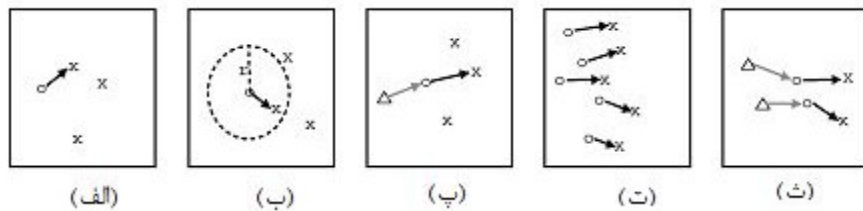
^۱ Contour

هدف روش‌های قطعی به حداقل رساندن تابع هزینه و مدل نمودن ردیابی به صورت یک مساله ی بهینه سازی می‌باشد.

هزینه ی تشابه عموماً توسط ترکیبی از محدودیت‌های زیر به دست می‌آید:

- مجاورت: فرض می‌نماید که مکان یک شی در دو فریم متوالی نمی‌تواند 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 شباهت‌های موجود را به یک همسایگی دایره شکل محدود می‌نماید. (شکل ۱-۳ ب)
- تغییرات جزئی سرعت: این محدودیت فرض می‌نماید که سرعت و جهت یک شی نمی‌تواند 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 می‌باشد. این شرط برای تشخیص وابستگی بین اشیائی که با استفاده از چندین نقطه نمایش داده می‌شوند مورد استفاده قرار می‌گیرد. (شکل ۱-۳ ت)
- ثبات شکل^۱ (انعطاف ناپذیری): این اصل اشیاء را در یک فضای سه بعدی در نظر می‌گیرد، 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)

^۱ Rigidity



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

۲-۳-۳ ردیابی بر اساس روش های پیش بینی^۱

در این گونه از روشها حالت شیء در گذشته ثبت می شود و بر اساس آن حالت جدید در قاب جاری پیش بینی می شود. سپس بر اساس اینکه تا چه حد این پیش بینی با شواهد همخوانی داشته باشد فرضیات جدید تولید می گردد. مهمترین و اولین روش در این گروه روش فیلتر کالمن است که در ادامه بیشتر توضیح داده خواهد شد. در این روش مشاهدات دارای توزیع گاوسی فرض شده که همیشه صادق نمی باشد. برای حل این مشکل روش های دیگری از جمله کالمن تعمیم یافته، *UKF* و تئوری فرضیات چندگانه (*MHT*) ارائه شده است که ما روش آخر را توضیح خواهیم داد.

از آنجاییکه داده های دریافت شده از سنسورها کاملاً دقیق نبوده و دارای نویز می باشند، روش های قطعی، همواره با مشکلاتی روبرو هستند. اما روش های بر مبنای پیش بینی، علاوه بر داده های اندازه گیری شده، عدم قطعیت را نیز در روال تشخیص و پیش بینی اجسام متحرک، در فریم های متوالی مد نظر قرار می دهند. و عموماً از یک فضای حالت برای مدل کردن ویژگی های شیء مانند، مکان، سرعت و شتاب استفاده می نمایند. حالات اشیاء در تصویر، بر اساس حالات تشخیص داده شده قبلی و معادلات تغییر حالت تخمین زده می شود. تخمین بدست آمده با مشاهدات در تصویر واقعی مقایسه شده و فضای حالت بروز رسانی می گردد. از میان تخمین زنده ها به فیلتر کالمن و ذره های می توان اشاره نمود، که در ادامه با آنها آشنا می شویم.

¹ Prediction

هدف از مدل احتمالی، مشخص کردن توزیع احتمال پسین مکان یک خودرو یا حرکت آن بر اساس اندازه گیری های به دست آمده از تصویر می باشد.

اگر x_t نشان دهنده ی وضعیت^۱ یک خودرو 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 مساله ی محاسبه ی توزیع احتمال در دنباله ی $x_{1:t} = (x_1, \dots, x_t)$ مشروط به دریافت دنباله ی مشاهدات $z_{1:t} = (z_1, \dots, z_t)$ یعنی $p(x_{1:t}|z_{1:t})$ تعریف می گردد. بنابر قانون بیز خواهیم داشت :

$$p(x_{1:t}|z_{1:t}) = \frac{p(z_{1:t}|x_{1:t})p(x_{1:t})}{p(z_{1:t})} \quad (1-3)$$

در رابطه ی بالا $p(z_{1:t}|x_{1:t})$ شباهت^۲ نامیده می شود

۳ -۴ - فیلترینگ تصادفی^۳

فیلترینگ فرایند استخراج اطلاعات در مورد کمیتی مورد نظر در لحظه ی t با در دست داشتن اطلاعات اندازه گیری^۴ تا لحظه ی t و همچنین شامل لحظه ی t می باشد. [۴۱]

پیش بینی^۵ شکل پیشین^۶ فرایند تخمین است که هدف از آن پیش بینی توزیع کمیت مورد نظر در زمان $t + \tau$ و $(\tau > 0)$ توسط اطلاعات اندازه گیری تا لحظه ی t می باشد. پیش بینی، یک ناحیه جدید 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 خودرو باشد.

^۱ State

^۲ likelihood

^۳ Stochastic Filtering Problem

^۴ Measured Data

^۵ Prediction

^۶ a prior

به روز رسانی^۱ شکل پسین^۲ 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 اندازه‌گیری در لحظه ی t انجام می‌شود.

مسئله‌ی فیلترینگ تصادفی در فضای حالت دینامیکی به شکل زیر در نظر گرفته می‌شود:

$$x(t) = f(t, x_t, v_t) \quad (۲-۳)$$

$$z(t) = h(t, x_t, n_t) \quad (۳-۳)$$

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

به دلیل اینکه فیلتر کردن برای ردیابی وسایل نقلیه در حوزه‌ی زمان گسسته انجام می‌گیرد. معادلات (۲-۳) و (۳-۳) را می‌توان به شکل زیر بازنویسی نمود:

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

$$z_k = h_x(x_k, n_k) \quad (۵-۳)$$

v_{k-1} و n_k نویز سفید با پارامترهای معلوم در حوزه‌ی زمان می‌باشند. معادله سیستم در (۴-۳)، معرف تابع توزیع احتمال گذر حالت^۵ به صورت $p(x_k | x_{k-1})$ و معادله‌ی اندازه‌گیری (۵-۳) معرف تابع

^۱ Updating

A posteriori^۲

^۲ Process Noise

^۴ Measurement Noise

^۵ State Transition Probability Density Function

شباهت به صورت $p(z_k|x_k)$ می‌باشد. هدف از فیلتر کردن، تخمین بهینه‌ی متغیر حالت در زمان با 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 $p(z_k|x_k)$ می‌باشد. مقدار تخمین‌زده شده به صورت $p(x_k|z_{0:k})$ یا $p(x_{0:k}|z_{0:k})$ خواهد بود که $x_{0,k}$ بردار حالت هدف از ابتدا تا لحظه‌ی جاری است.

۳-۴-۱ فیلترینگ بازگشتی بیزین غیر خطی^۱

قاعده بیز^۲ یک ابزار ریاضی برای استنتاج علت‌ها از یک سری مشاهدات می‌باشد. در ردیابی وسایل نقلیه 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 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^۳ و بروز رسانی^۴ 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 مرحله‌ی پیش‌بینی شامل استفاده از مدل سیستم (۳-۵) (z_n) جهت به دست آوردن تابع توزیع پیشین^۵ حالت در لحظه‌ی k با استفاده از معادله‌ی Chapman-Komogorov می‌باشد:

¹ Recursive Nonlinear Bayesian Filtering

² Bayes Rule

³ Prediction

⁴ Update

⁵ A posteriori Estimate

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

در بدست آوردن رابطه‌ی (۶-۳) 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 (۴-۳) و با در دست داشتن پارامترهای معلوم برای توزیع v_{k-1} قابل محاسبه است. بنابر تئوری بیزین خواهیم داشت:

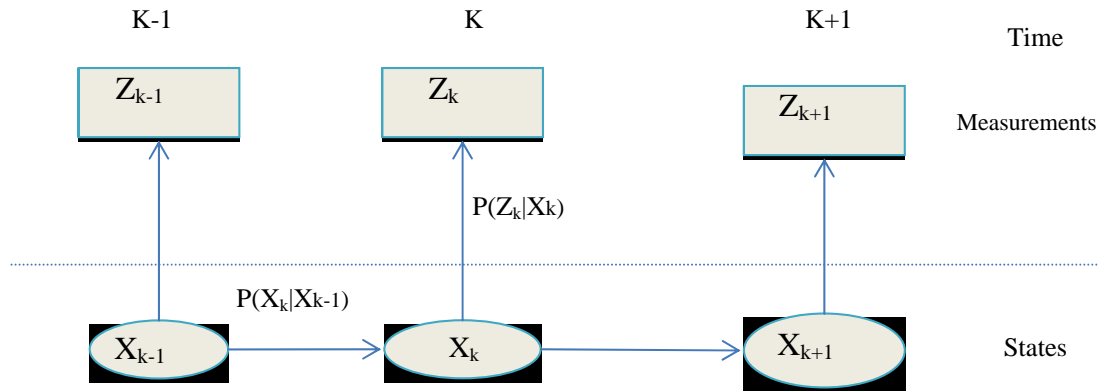
$$p(x_k | z_{1:k}) = \frac{p(z_k | x_k) p(x_k | z_{1:k-1})}{p(z_k | z_{1:k-1})} \quad (7-3)$$

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 (۸-۳) وابسته به توزیع به دست آمده از مرحله‌ی پیش بینی و همچنین تابع شباهت است که توسط مدل اندازه‌گیری (۵-۳) و با در دست داشتن پارامترهای توزیع n_k قابل محاسبه است. مراحل پیش بینی و بروز رسانی در شکل (۲-۳) نشان داده شده است.

معادلات بازگشتی (۶-۳) و (۷-۳) 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 شده اند که بسته به ویژگی توزیع حالت هدف و نوع کاربرد مورد نظر قابل استفاده‌اند.



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

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

$$X_t = f'(X_{t-1}) + W_t \quad (9-3)$$

که W_t نویز سفید می باشد. ارتباط بین مقدار اندازه گیری شده و بردار حالت توسط رابطه زیر محاسبه می گردد:

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

که N_t نیز نویز سفید جداگانه ای است .

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

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

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:

$$p(X_t | Z_{1:t-1}) \quad (۱۲-۳)$$

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:

$$p(X_t | Z_{1:t}) \quad (۱۳-۳)$$

۳ ۵ - کالمن:

فیلتر کالمن یک ابزار ریاضی قدرتمند برای تخمین داده‌ها و ردیابی اشیاء می‌باشد. [۴۲] 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 مانور جسم ثابت نبوده و در بازه‌های زمانی مختلف به صورت بهینه محاسبه می‌شود. بنابراین این فیلتر از دقت بالایی برخوردار می‌باشد. در حال حاضر ترکیبهای گسترده‌ای از فیلترهای کالمن بسط یافته 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 می‌تواند فقط با دریافت بعضی از این متغیرها سایر آنها را تخمین بزند. شاید مهمترین کاربرد فیلتر کالمن،

¹ Prediction

² Correction

³ Like Hood

⁴ Posterior

⁵ Gain

تخمین موقعیت شیء، هنگام خروج شیء از دید دوربین می‌باشد. در این موارد فیلتر کالمن بر اساس مشاهدات دریافت شده در زمان های قبل، قادر به تخمین مکان تقریبی شیء می‌باشد. فیلتر کالمن در زمانی 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 قله ای^۱ باشد. فیلتر کالمن توانایی پیش بینی در مواردی از قبیل وجود همپوشانی بین اشیاء و دینامیک پیچیده و غیر خطی اشیاء (توزیع احتمال پارامترهای حالت، چند قله ای می‌باشد)، را ندارد. همچنین مسائلی که دارای ماهیت خطی و گاوسی نیستند یا نویز آنها از نوع اضافه شونده گاوسی نباشد با فیلتر کالمن نتایج مطلوبی را نخواهند داشت.

• تعریف ریاضی فیلتر کالمن

در فیلتر کالمن فرض بر این است که تابع توزیع پسین برای متغیر حالت در هر لحظه به شکل یک توزیع گاوسی است 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:

- n_k و v_{k-1} 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 متغیرهایشان باشند.

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

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

$$z_k = h_k x_k + n_k \quad (۱۵-۳)$$

^۱ Uni-Modal

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:

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

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:

$$Z_k = H(X_k, W_k) \quad (۱۷-۳)$$

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

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

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

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.

تعریف: $i \geq j$ و $(X_i | j)$ ، را تخمین حالت X_i می نامیم. با داشتن تخمین حالت $(X_k | k)$ و ورودی

سیستم U_k می توانیم حالت $(X_{k+1} | k)$ را با استفاده از فرمول زیر تخمین بزنیم:

^۱ Recursive State Estimation

^۲ stochastic

^۳ Iterative

$$X_{k+1|k} = E\left[f(X_k, u_k, V_k) | Z^k\right] \quad (20-3)$$

اگر برای رابطه (3-17) از نتایج روابط بالا استفاده شود، تخمین توزیع پیشین از مقدار اندازه گیری سیستم با رابطه ی (3-21) برابر خواهد شد.

$$Z_{k+1|k} = h(X_{k+1|k}, 0) \quad (21-3)$$

با مقایسه مقدار واقعی بردار Z_{k+1} و مقدار تخمین زده شده ی $(Z_{k+1|k})$ می توان به مقدار خطای اندازه گیری دست یافت. با اضافه کردن این خطا به مقدار تخمین حالت پیشین، می توان تخمین بهتری از مقدار واقعی داشت که به آن تخمین حالت پسین¹ دست یافت:

$$X_{k+1|k+1} = X_{k+1|k} + K_{k+1}(Z_{k+1} - Z_{k+1|k}) \quad (22-3)$$

k_{k+1} بهره ی² کالمن می باشد که برای مینیمم کردن کواریانس خطای حالت پسین استفاده می شود. 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 کواریانس). مقداری که برای k بدست می آید مقدار بهینه ای برای رابطه ی (3-22) خواهد بود.

خطای تخمین حالت پیشین و پسین به صورت زیر محاسبه می شوند:

$$e_{k+1|k} = X_{k+1} - X_{k+1|k} \quad (23-3)$$

$$e_{k+1|k+1} = X_{k+1} - X_{k+1|k+1} \quad (24-3)$$

با توجه به مدل خطی در رابطه ی (3-17) تخمین حالت پیشین بدون نویز از رابطه زیر محاسبه می شود

¹ A posteriori
² Gain

$$X_{k+1|k} = F_k X_{k|k} + u_k \quad (25-3)$$

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

$$e_{k+1|k} = F_k e_{k|k} + V_k \quad (26-3)$$

مقادیر z ، نیز از روابط زیر بدست می آید:

$$Z_{k+1|k} = H_{k+1} X_{k+1|k} \quad (27-3)$$

$$Z_{k+1} = H_{k+1} X_{k+1} + W_{k+1} \quad (28-3)$$

با استفاده از رابطه ی (24-3) می توان یک رابطه بازگشتی برای به دست آوردن خطای تخمین

حالت پسین بدست آورد:

$$e_{k+1|k+1} = e_{k+1|k} - k_{k+1} (H_{k+1} e_{k+1|k} + W_{k+1}) \quad (29-3)$$

مقدار کوواریانس خطا با توان دوم خطا بدست می آید:

$$P_{k+1|k} = E[e_{k+1|k} e_{k+1|k}^T] \quad (30-3)$$

$$P_{k+1|k+1} = E[e_{k+1|k+1} e_{k+1|k+1}^T] \quad (31-3)$$

با استفاده از روابط (26-3) و (28-3) می توان به رابطه ساده تری برای $(P_{k+1|k})$ دست یافت:

$$P_{k+1|k} = F_k P_{k|k} F_k^T + Q_k \quad (32-3)$$

با جایگذاری رابطه (26-3) در رابطه (29-3) رابطه کوواریانس خطای پسین به صورت زیر خواهد

بود:

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k} H^T K^T - K H P_{k+1|k} + K (H P_{k+1|k} H^T + R) K^T \quad (3-33)$$

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).

$$K = P_{k+1|k} H^T (H P_{k+1|k} H^T + R)^{-1} \quad (3-34)$$

روش ارائه شده برای فیلتر کالمن هنگام صادق بودن تمامی فرضیات در نظر گرفته شده، بهینه ترین حالت برای مسئله ی تخمین حالت می‌باشد. به این معنی که در صورت خطی و گاوسی بودن توزیع حالت هیچ الگوریتمی منجر به جواب بهتر نخواهد شد. در برخی از کاربرد های عملی نمی‌توان تمام فرض های لازم برای فیلتر کالمن را فراهم نمود. در این گونه مواقع می‌توان از الگوریتم های زیر اپتیمم^۱ استفاده نمود. یکی از این الگوریتم ها الگوریتم کالمن توسعه یافته^۲ [۴۳] است.

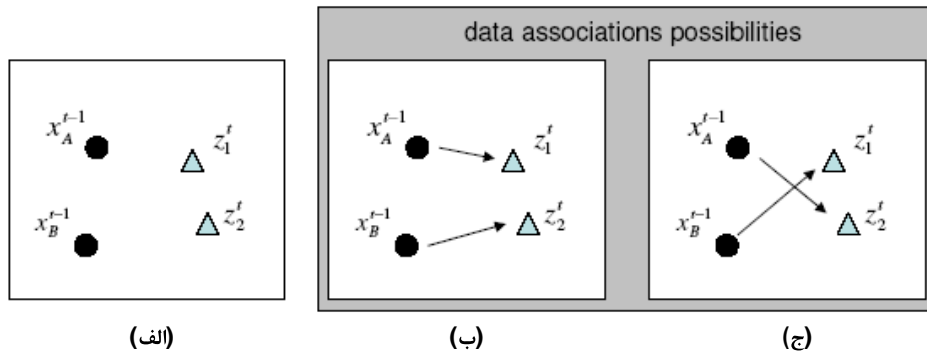
۳-۶ - ردیابی اهداف چندگانه

مسئله ی ردیابی اهداف چند گانه^۳ (MTT) را علیرغم شباهت زیاد آن به ردیابی اهداف تکی، نمی‌توان 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 وجود هدف B نیز در مکان x_B^{t-1} تثبیت شده است.

¹ Sub optimal Algorithm

² Extended Kalman Filter

³ Multiple-Target Tracking



شکل ۳-۳: مثال ارتباط اطلاعات. الف) دو هدف (دایره‌ها) به همراه دو داده‌ی مشاهده شده (مثلث‌ها)، جهت ایجاد ارتباط. ب) حالت ممکن اول در مسئله‌ی ارتباط اطلاعات. ج) حالت ممکن دوم برای مسئله‌ی ارتباط اطلاعات.

در شکل (۳-۴ الف) اطلاعات Z_1^t و Z_2^t در زمان t توسط سیستم مشاهده می‌شوند. ابهام موجود در مسئله‌ی 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 داده‌ی اندازه‌گیری تولید کرده و هیچ آشکار سازی نادرست^۱ در سیستم صورت نگرفته است (عدم وجود پارازیت^۲ در محیط) دو حالت برای ارتباط اطلاعات مشاهده و مسیر حرکت^۳ می‌توان در نظر گرفت:

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

- Z_1^t متناظر با A و Z_2^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 آشکار سازی 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 اندازه‌گیری توسط هر هدف

¹ False Alarm

² Clutter

³ Track

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.

۳-۶-۱ 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 می‌توان به صورت مسئله فیلترینگ بیزین متوالی^۱ با مشاهدات نويز 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:

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

با توجه به رابطه فوق، فضای حالت کلی برابر اجتماع حالاتی است که هدفی موجود نباشد (\emptyset) یک هدف داشته باشیم (x) و یا حالت‌های متشکل از تعداد محدودی هدف (x^N) باشد. مدل گذر از حالت چند هدفه‌ی s^{t-1} در لحظه‌ی $t-1$ به حالت s^t در لحظه‌ی t را می‌توان توسط احتمال گذر $p(s^t | s^{t-1})$ بیان کرد که مشابه معادله‌ی سیستم (دینامیک هدف) در ردیابی تک هدفه است؛ با این تفاوت که لازم است چگونگی ورود و خروج اهداف نیز مدل شوند. در ردیابی چند گانه اطلاعات اندازه‌گیری مجموعه‌ای از مشاهدات تولید شده توسط اهداف، و مشاهدات نادرست حاصل از وجود خطا در سنسورها می‌باشد. بنابراین اگر Z فضای اندازه‌گیری^۲ مربوط به یک هدف باشد فضای کلی اندازه‌گیری در حالت چند هدفه اجتماع تمام مجموعه‌های Z خواهد بود که آنرا با M نشان

¹ Sequential Bayesian Filtering Problem

² Measurement Space

می‌دهیم. در اینصورت مدل اندازه‌گیری توسط یک تابع شباهت به صورت $P(m|Z')$ ارائه خواهد شد که در آن m' مجموعه مشاهدات می‌باشد که مقادیر خود را از M خواهد گرفت. در مثال دو هدفه‌ی نشان داده شده در شکل (۳-۴) مجموعه مشاهدات به صورت $m' \in \{Z'_1, Z'_2\}$ می‌باشد. مسئله ارتباط اطلاعات به دلیل معلوم نبودن ترتیب اطلاعات موجود در مجموعه مشاهدات حاصل می‌شود.

از لحاظ تئوری می‌توان با اعمال مستقیم روش فیلترینگ بیزین بازگشتی استاندارد به مسئله عمومی ردیابی اهداف چندگانه مطرح شده در بخش قبلی، به توزیع فیلتر شده‌ی مطلوب $p(s|m^0, \dots, m')$ دست یافت. با این حال محاسبه توزیع فیلتر شده‌ی 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 سیستم‌های ردیابی اهداف چندگانه مدیریت پیچیدگی و حجم محاسباتی در سیستم با در نظر داشتن عملکرد مناسب برای عمل ردیابی است.

روش‌های زیادی برای حل مشکل ارتباط اطلاعات در ردیابی اهداف چندگانه وجود دارد. یکی از این روش‌ها MHT می‌باشد که تمام حالات موجود برای مسیر موجود بین مشاهدات و اهداف را در نظر می‌گیرد. این باعث می‌شود 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 روش‌های دیگر حل این مشکل فیلتر استاندارد نزدیکترین همسایگی^۱ می‌باشد که هر کدام از اهداف را به نزدیکترین مشاهده در فضای هدف ارتباط می‌دهد. اما این روش باعث از بین رفتن یک سری از فرضیات امکان پذیر و شدنی می‌شود. فیلتر $JPDAF$ یکی از روش‌های مناسبی است که برای حل مسئله‌ی ارتباط اطلاعات استفاده می‌شود.

¹ Nearest Neighbour Standard Filter

۳-۶-۲ ردیابی فرضیات چندگانه^۱(MHT)

این روش تمام ارتباطات^۲ (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. [۴۴] 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.

۳-۶-۳ الگوریتم JPDAF^۳

این روش شامل بروز رسانی هر یک از حالات مسیر توسط ترکیبات وزن دار تمامی مشاهدات می باشد که توسط فرتمن، شالم و شف^۴ [۴۵] 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 $p(s | m^0, \dots, m')$ 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^۵ با هر مسیر

¹ Multiple Hypothesis Tracking

² Hypothesis

³ Joint Probability Data Association Filter

⁴ Fortman , Shalom , Scheffe

⁵ Soft Data Association

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 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:

- 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:

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

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

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 مسیره‌ها در رابطه (۳۶-۳) 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 فرضیه‌ها حالت مسیره‌های A, B بر پایه قانون بی‌زین و با استفاده از مشاهده مربوط شده با آن به روز خواهد شد. در نتیجه، توزیع مسیره 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 نرمالیزاسی می باشند. حالت مسیره با مشاهدات مختلف تحت فرضیه های متفاوت به روز می شود؛ بنابراین الگوریتم MHT براساس هر یک از فرضیه ها حالت مسیره جداگانه ای را حفظ خواهد کرد. در صورتی که الگوریتم $PDFA$ فرضیه های چندگانه را با مخلوط کردن توزیع احتمال مسیره‌های یکسان براساس تمام فرضیه ها ترکیب کرده و به فرضیه ی واحدی دست خواهد یافت که مفهوم تصمیم گیری نرم برای محاسبه ارتباطات در هر مسیره می باشد.

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

۳-۶-۴ فیلتر ذره‌ای

علاوه بر فیلتر کالمن و مشتقات آن، روش‌های دیگری نیز برای ردیابی اهداف ارائه شده اند. یکی از مشهورترین این روش‌ها فیلتر ذره‌ای می‌باشد. ایده اصلی فیلتر ذره‌ای اعمال بازگشتی فیلتر بیزین بر روی مجموعه‌ی نمونه‌هاست. [۴۶] این فیلتر به سیستم‌های خطی محدود نبوده و لزومی به محدود کردن نویز به نوع گاوسی اضافه شونده نیز وجود ندارد. روش فیلتر ذره‌ای اخیراً در محیط‌های غیر گاوسی برتری‌هایی نسبت به روش فیلتر کالمن نشان داده است. اما پیچیدگی‌های محاسباتی بالای روش‌های مبتنی بر فیلتر ذره‌ای مانع از کاربرد آنها در مسائل بزرگ نظیر ردیابی تعداد زیاد اشیاء متحرک می‌شود. به عنوان مثالی دیگر می‌توان به الگوریتم چگالش^۱ اشاره نمود [۴۷]. این الگوریتم توسعه‌ای از فیلتر ذره‌ای می‌باشد که قادر به ردیابی توزیع‌های چند قله‌ای نیز می‌باشد و با عناوینی نظیر فیلتر بوت استرپ بیزی^۲ و فیلتر مونت کارلو^۳ نیز شناخته می‌شود. از آنجائیکه این الگوریتم یک روش غیر پارامتریک است به تعداد نسبتاً زیادی نمونه جهت تضمین یک تخمین بیشترین احتمال^۴ صحیح از حالت فعلی نیاز دارد.

این فیلتر برای تقریب تابع چگالی احتمال پسین از مجموعه‌ای از نمونه‌های وزن‌دار ذرات استفاده می‌کند. در این فیلتر بر اساس مدل حرکتی انتخاب شده برای هدف عملیات انتشار تکراری ذرات صورت می‌گیرد و بر اساس مجموعه‌ای از مشاهدات، ذرات وزن دهی می‌شوند.

فیلتر ذره‌ای با استفاده از مجموعه‌ای از N_i مجموعه وزن‌دار تصادفی، برای هر فریم زمانی به تخمین چگالی احتمال پیشین می‌پردازد:

$$p(X_k | Z_{1:k}) \approx \sum_{i=1}^N w_k^{(i)} \delta(X_k - X_k^{(i)}) \quad (۴۳-۳)$$

^۱ Condensation

^۲ Bayesian Bootstrap Filter

^۳ Monte Carlo Filter

^۴ Maximum Likelihood Estimation(MLE)

که مجموع وزن‌های اسکالر $w_k^{1, \dots, N}$ واحد است و تابع $\delta(\cdot)$ دلتای دیراک می‌باشد. در ادامه بر اساس یک توزیع احتمالی ناشی از وزن‌های ذرات، عملیات نمونه برداری مجدد^۱ انجام می‌شود. در نهایت موقعیت هدف با یک میانگین‌گیری از موقعیت ذرات بدست می‌آید. در عمل با داشتن یک مدل احتمالی بر اساس بیشترین احتمال، بعد از چند تکرار ذرات در نواحی که احتمال بالاتری دارند، متمرکز می‌شوند.

یک روش ساده برای پیاده‌سازی فیلتر ذره‌ای، روش SIR^۲ می‌باشد. در این الگوریتم ذرات بر اساس یک مدل حرکتی انتشار می‌یابند؛ این مدل بر اساس دانشی از حرکت وسایل نقلیه انتخاب می‌شود. و توسط آن می‌توان نشان داد که واریانس وزن‌ها در طول زمان افزایش می‌یابد، که این امر باعث کاهش دقت الگوریتم می‌گردد؛ این مساله به انحطاط^۳ معروف است و برای رفع این مشکل از نمونه برداری مجدد^۴ استفاده می‌شود؛ که در آن ذرات دارای وزن‌های کوچک با ذراتی با وزن بیشتر جایگزین می‌شوند. موقعیت وسیله نقلیه متحرک در فریم زمانی k با رابطه زیر بدست می‌آید:

$$X_k \approx \sum_{i=1}^N w_k^{(i)} X_k^{(i)} \quad (۴۴-۳)$$

بعد از تخمین توزیع چگالی، نقاط تخمینی برای نمایش موقعیت خودرو انتخاب می‌شوند و یک فیلتر میانگین‌گیر در ادامه به کار گرفته می‌شود.

$$X_k^{ME} = \text{median}\{x_k^1, x_k^2, \dots, x_k^N\} \quad (۴۵-۳)$$

فرض: در فریم زمانی $k-1$ مجموعه‌ای از پارتیکل‌ها $X_{k-1}^{(i)}$ و وزن‌ها $w_{k-1}^{(i)}$ به صورت گسسته‌ای $p(X_{k-1} | Z_{1:k-1})$ را به نمایش می‌گذارند.

^۱ Resampling

^۲ Sampling Importance Resampling

^۳ degeneracy

^۴ Resampling

تکرار: مشاهده Z_k از فریم زمانی k و بروز رسانی ذرات با توجه به روابط زیر:

۱- پیش‌بینی: انتشار ذرات از طریق معادله انتقالی

۲- بروز رسانی: تخصیص دادن یک وزن احتمالی به هر ذره و نرمالیزه کردن وزن‌ها.

For $j=1$ to N do

$$\hat{X}_k^j \approx p(X_k | X_{k-1}^j)$$

$$\text{Calculate } w_k^j = p(Z_k | X_k^j)$$

END

$$w_k^{1:N} / \sum_{j=1}^N w_k^j$$

۳- نمونه برداری مجدد: بدست آوردن N نمونه جدید $X_k^{(i)}$ از پارتیکل‌های موجود $\{\hat{X}_k^j\}_{j=1}^N$ بر

طبق وزن‌های آن‌ها، سپس ریست کردن وزن‌ها $w_k^{(i)} = 1/N$

$$c_1 = 0$$

For $j=2$ to N do

$$\text{Calculate CDF: } c_j = c_{j-1} + w_k^j$$

END

$$j=1$$

$$\text{Starting point: } u_1 = U(0, N_j^{-1})$$

For $n=1$ to N do

$$u_n = u_1 + (n-1)N_j^{-1}$$

While $u_n > c_j$ do

$$j=j+1$$

END

$$X_k^{(n)} = X_k^j, w_n^{(i)} = 1/N$$

نتیجه: مجموعه $\{X_k^{(i)}, w_k^{(i)}\}_{i=1}^N$ توزیع چگالی $p(X_k | Z_{1:k})$ را تخمین می‌زند.

الگوریتم ۱: الگوریتم نمونه‌برداری مجدد برای فیلتر ذره‌ای

۳-۶- نتیجه گیری فصل

در این فصل ابتدا مرور مختصری بر کارهایی که در زمینه ردیابی صورت گرفته است داشتیم، سپس روش های ردیابی وسایل نقلیه که خود به دو دسته اصلی روش های مبتنی بر قطعیت و روش های مبتنی بر پیش بینی است مورد بررسی قرار گرفت. در ادامه فیلتر کالمن و الگوریتم MHT، بعنوان روش استفاده شده در این پایان نامه مورد بررسی دقیق تر قرار گرفت. همانگونه که در این فصل به آن اشاره شد، علت انتخاب فیلتر کالمن این است که در تخمین و ردیابی اجسامی که حرکتی خطی دارند، بهینه ترین پاسخ را می دهد. از طرفی فیلتر ذره ای با افزایش تعداد وسایل نقلیه متحرک، بسیار کند عمل می نماید و از این رو، در یک سیستم بلادرنگ استفاده چندانی ندارد. در نتیجه با توجه به تعداد بالای خودروهای موجود در تقاطع ها، و اینکه مسیر حرکتی خودروها تقریباً خطی می باشد و پیچیدگی بسیار بالایی در حرکت آنها مشاهده نمی شود، از روش فیلتر کالمن به همراه تئوری ردیابی فرضیات چندگانه برای ردیابی همزمان تعداد بالایی از وسایل نقلیه متحرک استفاده نمودیم.

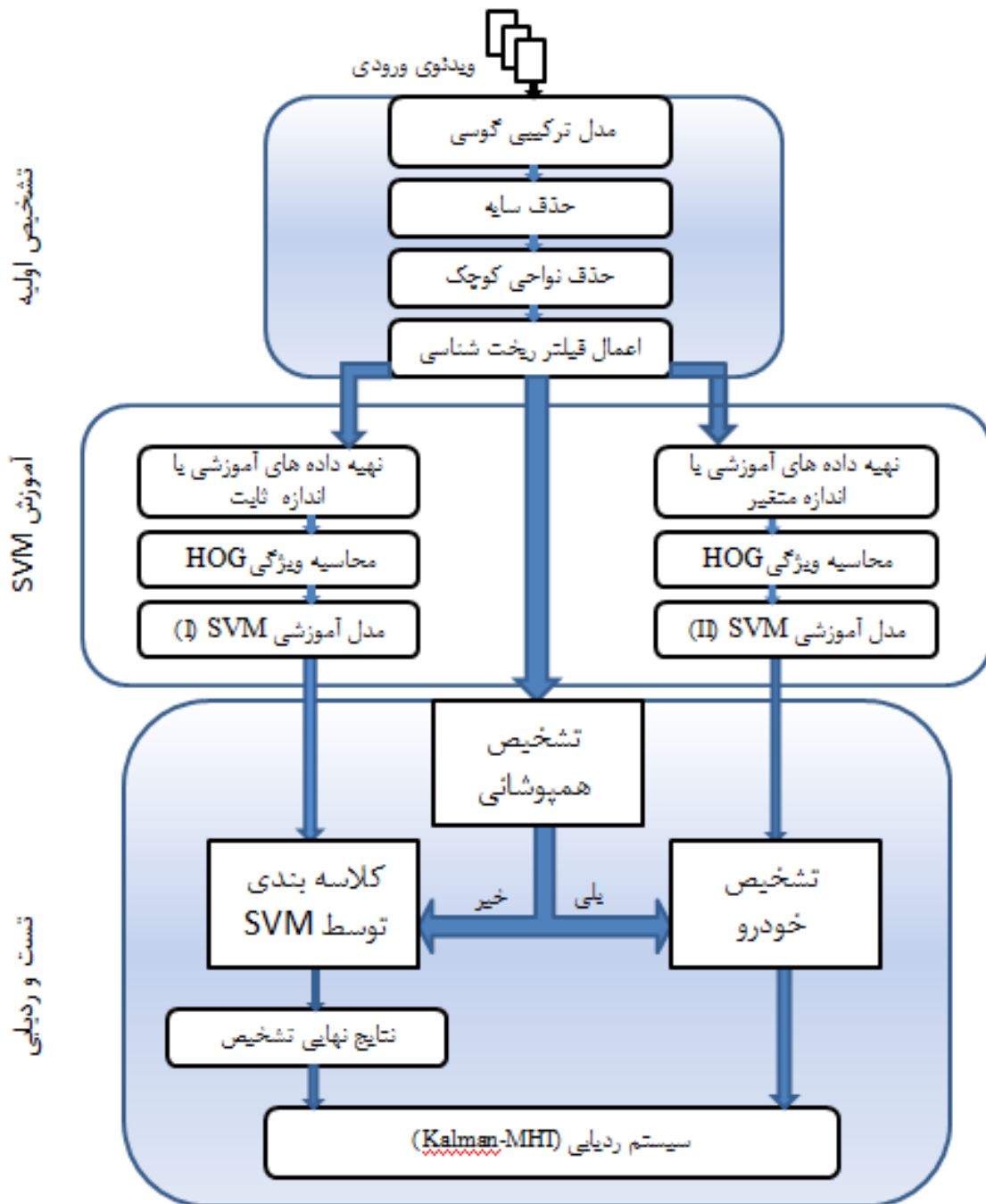
فصل چهارم

الگوریتم‌های پیشنهادی برای تشخیص وردیابی

۱-۴ - مقدمه

در فصل‌های قبلی مروری بر کلیه روش‌های تشخیص و ردیابی وسایل نقلیه متحرک داشتیم، همچنین مشکلات این فرآیند در تقاطع تا حدودی مورد بحث قرار گرفت. در این فصل این مشکلات، با جزئیات بیشتری مطرح خواهد شد و سپس الگوریتم‌های پیشنهادی برای حل آن‌ها را مورد بررسی قرار خواهیم داد.

در شکل ۱-۴ شمای کلی سیستم پیشنهادی برای تشخیص و ردیابی وسایل نقلیه متحرک نمایش داده شده است.



شکل ۴-۱: شمای کلی الگوریتم پیشنهادی برای تشخیص و ردیابی وسایل نقلیه متحرک

همانطور که در فصل دوم آورده شده بود، روش مدل ترکیبی گوسی به عنوان یکی از مناسب‌ترین روش‌ها در تشخیص اشیاء متحرک به شمار می‌آید. خروجی این مرحله علاوه بر اشیاء متحرک، شامل نویزهای مختلف و سایه می‌باشد، که در اثر شرایط نامناسب آب و هوایی، کیفیت پایین ویدئوهای ترافیکی و ... ایجاد می‌گردند و در بخش‌های بعدی مورد تحلیل قرار می‌گیرند. به همین دلیل با

استفاده از فیلترهای مختلف، تلاش می‌کنیم که این نویزها را به حداقل برسانیم. خروجی مرحله تشخیص علاوه بر وسایل نقلیه متحرک، (در این تحقق هدف ردیابی انواع خودرو می‌باشد) شامل عابران پیاده، دوچرخه و موتور سیکلت می‌باشد، که در ادامه از آنها با عنوان غیر خودرو نام خواهیم برد.

برای حذف غیر خودرو نیازمند یک سیستم با یادگیری می‌باشیم که آن را با تصاویر مختلفی از خودروها که از زوایای مختلف و در شرایط جوی گوناگون تهیه شده است، آموزش دهیم. برای این منظور از ماشین بردار پشتیبان که به عنوان یکی از روش‌های طبقه بندی مطرح می‌باشد استفاده می‌نماییم. برای ورودی ماشین بردار پشتیبان از توصیف‌گر هیستوگرام گردیان جهت‌دار (HOG)¹ استفاده می‌نماییم. از آنجاییکه خودروها از زوایای دید مختلفی وارد تقاطع‌ها می‌شوند و همچنین از تصاویر دوربینی‌های ترافیکی که با فواصل مختلفی نسبت به تقاطع قرار دارند، استفاده می‌نماییم، نمی‌توان از یک ویژگی خاص مانند چرخ ماشین (برای تصاویر از کنار)، شیشه ماشین (تصاویر از جلو)، بخش مستطیلی فرم (برای تصاویر از عقب یا تصاویر هوایی)، برای تشخیص خودروها بهره گرفت. بنابراین نیازمند ویژگی‌های قدرتمندی برای تمایز خودرو از غیر خودرو می‌باشیم که نسبت به جهت و اندازه خودرو وابستگی نداشته باشد. به همین منظور از توصیف‌گر HOG به عنوان یکی از جدیدترین و قدرتمندترین توصیف‌گرها استفاده می‌نماییم.

علاوه بر غیر خودرو مشکلاتی از قبیل همپوشانی وسایل نقلیه، تشخیص ناقص یک خودرو، یکی شدن بلاب‌های خروجی از مرحله تشخیص، که مربوط به دو یا چندین وسیله نقلیه می‌باشد وجود دارد. برای رفع مشکلات فوق از یک سیستم تشخیص همپوشانی و دو مدل آموزشی SVM، همانطور که در شکل ۴-۱ دیده می‌شود، استفاده می‌نماییم. جزئیات مربوط به چگونگی عملکرد این قسمت در بخش‌های بعدی مورد بررسی قرار خواهد گرفت. خروجی این مرحله در شرایط مطلوب، وسایل نقلیه متحرک می‌باشد که عاری از مشکلات مطرح شده‌ی فوق می‌باشد. این خروجی به قسمت ردیابی که

¹ Histogram of Oriented Gradient

بر پایه فیلتر کالمن و تئوری ردیابی فرضیات چندگانه^۱ استوار است وارد خواهد شد. در ادامه این فصل هر بخش به طور مجزا مورد بررسی قرار می‌گیرد. و خروجی‌های مربوطه را نمایش خواهیم داد. در نهایت با استفاده از جداول مختلف به ارزیابی هر بخش می‌پردازیم.

۴-۲- تشخیص اولیه وسایل نقلیه متحرک

در این بخش نتایج پیاده‌سازی الگوریتم‌های مربوط به تشخیص اولیه وسایل نقلیه متحرک مورد بررسی قرار خواهد گرفت، همان‌طور که در بخش‌های قبلی آورده شد، ما از روش مدل ترکیبی گوسی برای حذف پس‌زمینه و جداسازی وسایل نقلیه متحرک از سایر اشیاء ثابت موجود در تصویر ویدئویی ورودی استفاده خواهیم نمود. در ادامه حذف نویزهای مختلف و حذف سایه از تصاویر خروجی پیاده‌سازی می‌شوند.

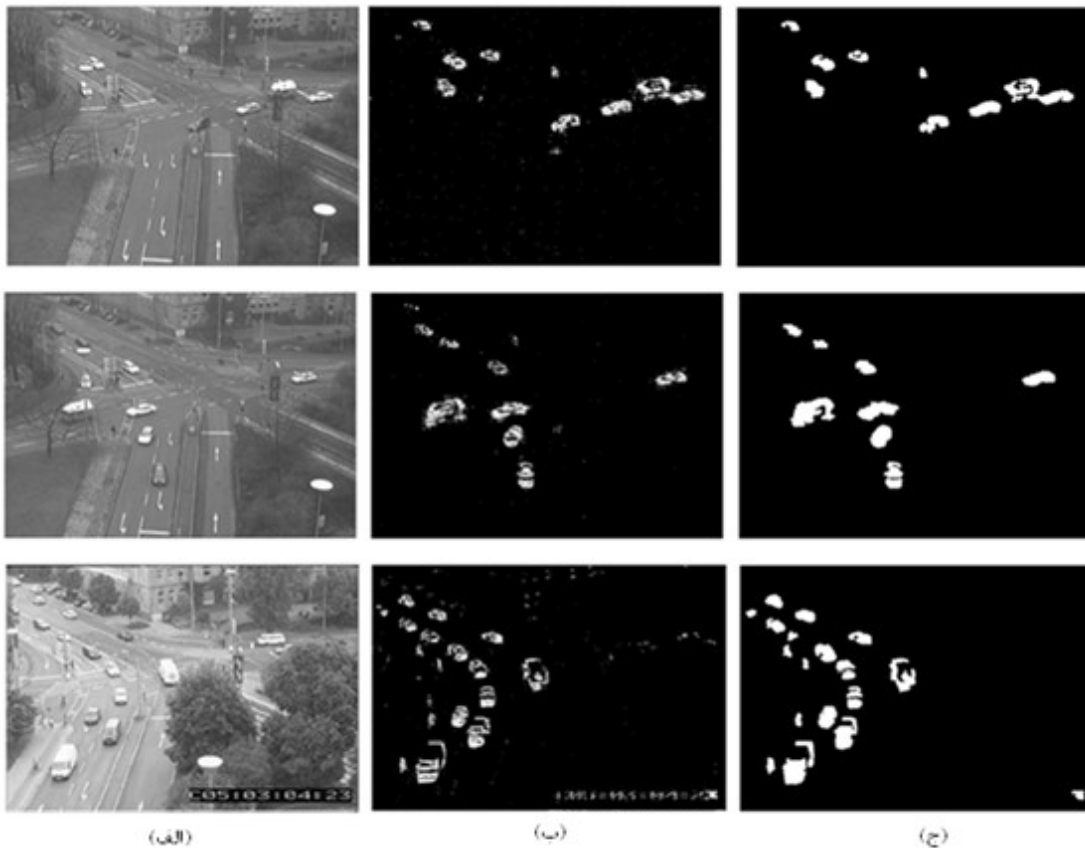
۴-۲-۱ مدل ترکیبی گوسی و حذف نویز

در این تحقیق، مدل ترکیبی گوسی به عنوان یکی از قوی‌ترین روش‌های تشخیص اشیاء متحرک مورد استفاده قرار می‌گیرد. این روش در فصل دوم معرفی و مورد بررسی قرار گرفت، در این بخش به پیاده‌سازی این روش، به همراه حذف نویزهای مختلف از آن می‌پردازیم. [۴۸]

وجود انواع نویزهای فلفل نمکی، وجود ذرات کوچک و سایر نویزهای ناشی از کیفیت پایین ویدئوهای ترافیکی ورودی، شرایط نامناسب جوی، برف و باران، مه غلیظ، به هم پیوستگی خودروهای نزدیک‌به‌هم، نیاز به استفاده از فیلترهای مختلف حذف نویز در خروجی مدل گوسی را ایجاد نموده است. در این تحقیق همان‌طور که در شکل ۴-۱ آورده شده است، پس از جداسازی اشیاء متحرک، حذف سایه انجام خواهد شد و در ادامه با استفاده از فیلترهای متوسط‌گیر تصاویر را تا حدودی هموار نموده تا اندکی از نویزهای فلفل - نمکی حذف شوند، سپس با استفاده از فیلترهای مناسب لکه‌های

^۱ Multiple Hypothesis Tracking (MHT)

کوچک را حذف نموده و در ادامه با استفاده از عملیات ریخت شناسی^۱، تا حد ممکن نویزها را از تصویر حذف می‌نماییم. در فرایند حذف نویز، نمی‌باید اطلاعات مفید وسایل نقلیه، حذف شوند، بنابراین در هموارسازی تصاویر خروجی به این مسأله توجه می‌کنیم. در شکل ۲-۴ نتایج خروجی این بخش روی تصاویر ویدئویی مختلف نمایش داده خواهد شد.

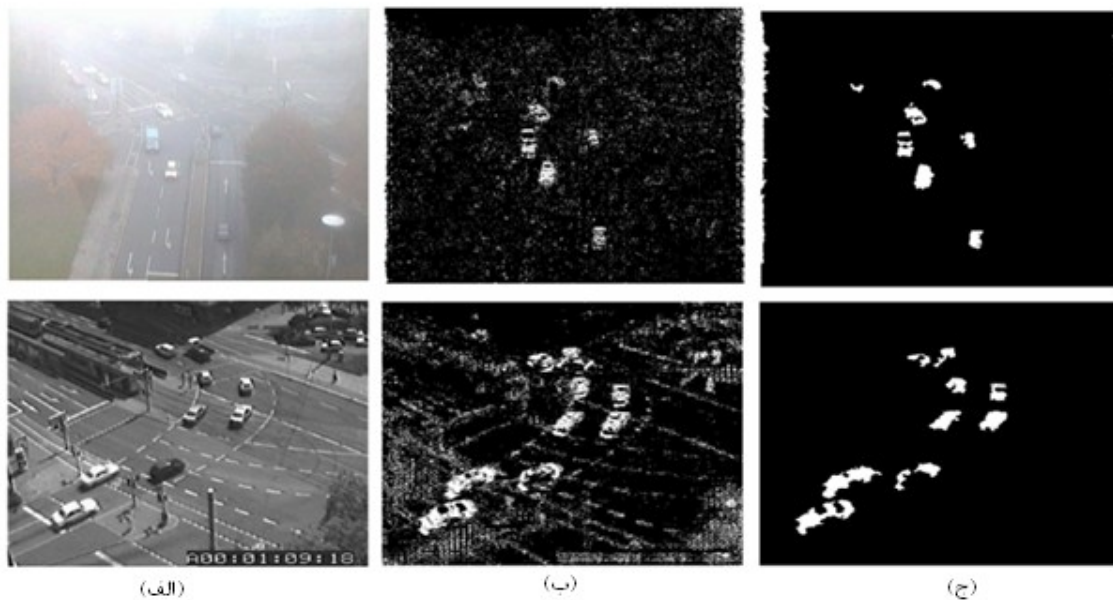


شکل ۲-۴: (الف) تصویر فریم‌های ورودی (ب) پس از مرحله حذف نویز پس زمینه GMM - (ج) پس از حذف نویز و حذف نواحی کوچک و همچنین اعمال فیلترهای مورفولوژی

همانطور که در تصاویر فوق دیده می‌شود، تصاویر خروجی نویزی، پس از اعمال فیلترهای مختلف حذف نویز، به خوبی هموار شده‌اند، البته تا حدودی ویژگی‌هایی مانند گوشه‌ها و لبه‌ها تحت تاثیر قرار گرفته‌اند. تصاویر فوق مربوط به ویدئوهای با شرایط معمولی جوی بوده‌اند، در شکل ۳-۴ تصویر

1- morphology

مربوط به شرایط جوی نامناسب (ردیف اول)، و ویدئوهایی با نویزهای بالا (ردیف دوم)، نمایش داده می‌شود:



شکل ۴-۳: (الف) تصویر فریم‌های ورودی در شرایط نویزی (ب) پس از مرحله حذف پس زمینه با استفاده از الگوریتم *GMM*. (ج) پس از حذف نویز و حذف نواحی کوچک و همچنین اعمال فیلترهای مورفولوژی

۴-۲-۲ حذف سایه از تصاویر مرحله تشخیص

تصاویری که در شرایط آفتابی گرفته شده‌اند معمولاً نیازمند اعمال فرایند حذف سایه می‌باشند، در این بخش در مورد سیستم ساده‌ی حذف سایه بحث و نتایج مربوطه نمایش داده خواهد شد.

در اغلب موارد، سایه‌ی اشیاء نیز به عنوان قسمتی از آن‌ها شناخته می‌شود؛ در حالی که سایه قسمتی از زمینه‌ی تصویر می‌باشد. بنابراین عملیات حذف سایه تأثیر بسزایی در افزایش کارایی سیستم ردیابی خواهد داشت. الگوریتم‌های متنوعی برای حذف سایه وجود دارد. به عنوان مثال در [۴۹] فرض شده است هنگامیکه سایه‌ی بر روی یک شیء می‌افتد، شدت سطح خاکستری آن شیء تغییر می‌نماید، در حالی که مقدار رنگ نرمال شده تغییر قابل توجهی نمی‌کند. به عبارت دیگر اگر $I(x,y)$ مقدار شدت سطح خاکستری را در نقطه (x,y) نشان دهد، خواهیم داشت:

$$\frac{R_s}{R_s + G_s + B_s} \cong \frac{R}{R + G + B}, \frac{G_s}{R_s + G_s + B_s} \cong \frac{G}{R + G + B}, \frac{B_s}{R_s + G_s + B_s} \cong \frac{B}{R + G + B} \quad (1-4)$$

$$I_s(x, y) = \alpha I(x, y) \quad (2-4)$$

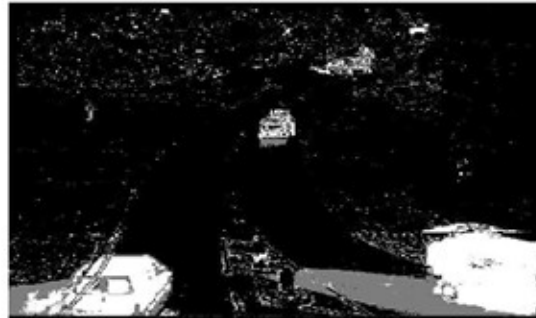
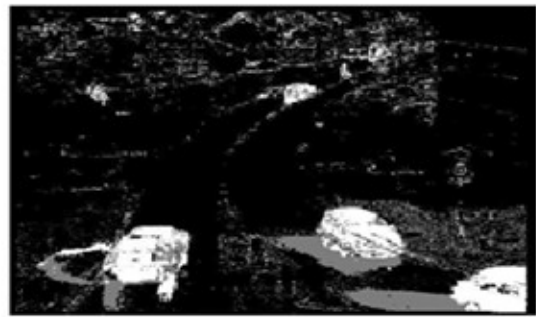
در روش دیگری که در مقاله‌ی [۵۰] معرفی شده است، میزان درخشندگی هر پیکسل به عنوان معیاری برای تشخیص سایه استفاده شده است. هنگام افتادن سایه بر روی یک سطح، مولفه‌ی درخشندگی بیشتر از مولفه‌هایی رنگی نور تحت تاثیر قرار می‌گیرند. اگر تصویر را به فضای رنگ HSV نگاشت نماییم، با استفاده از رابطه زیر می‌توان سایه را از تصویر جدا کرد.

$$shadow_{x,y}^k = \begin{cases} 1 & \text{if } D_H < \tau_h, |I_{x,y_s}^k - B_{x,y_s}^k| < \tau_s, v_a \leq \frac{I_{x,y_v}^k}{B_{x,y_v}^k} \leq v_b \\ 0 & \text{otherwise} \end{cases} \quad (3-4)$$

$$D_H = \min \left\{ |I_{x,y_h}^k - B_{x,y_h}^k|, 360 - |I_{x,y_h}^k - B_{x,y_h}^k| \right\} \quad (4-4)$$

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 (۳-۴) کنترل کرد.

شکل ۴-۴ نتیجه‌ی حاصل از تشخیص سایه توسط الگوریتم اول را نشان می‌دهد. در این تصاویر قسمت خاکستری ناحیه‌ای از خروجی مرحله‌ی تشخیص اشیاء است که به عنوان سایه شناخته شده است. همانگونه که مشاهده می‌شود حجم عمده‌ای از پیکسل‌ها، به عنوان سایه تشخیص داده شده‌اند.



(الف)

(ب)

شکل ۴-۴: (الف) تصاویر دو فریم ورودی (ب) پس از تشخیص اشیاء متحرک و تشخیص سایه

همان‌طور که در این بخش اشاره شد، عملیات تشخیص اولیه با استفاده از مدل ترکیبی گوسی و فیلترهای حذف نویز و حذف سایه مورد بررسی قرار گرفت. در صورتیکه ویدئوهایی با ترافیک سبک در نظر گرفته شود، می‌توان از این مرحله مستقیماً به مرحله ردیابی رفت. نمونه‌ای از یک ویدئو با تعداد بسیار کم خودرو در شکل ۴-۵ نشان داده شده است.



شکل ۴-۵: (الف) تصاویر خروجی از مرحله تفریق پس زمینه- GMM (ب تا ج) ردیابی در فریم های متوالی

همانطور که مشاهده می شود، در تصویر ویدئویی فوق وسایل نقلیه به راحتی با استفاده از سیستم ردیابی، در فریم های متوالی دنبال می شوند. رنگ قرمز مربوط به مرحله پیش بینی و رنگ سبز مرحله ی تثبیت را نشان می دهد. در این شکل خودرو ها با یکدیگر فاصله مناسبی دارند، تعداد خودروها کم است، شرایط جوی مناسب بوده، عابر پیاده و موتور سیکلت در تصاویر فوق دیده نمی شود، محل قرارگیری دوربین نسبت به خودروها مناسب می باشد و تصاویر با کیفیتی بالایی گرفته شده است. در حالیکه در اکثر مواقع در چهارراه و تقاطع ها، پیچیدگی ترافیکی بسیار بالا بوده و هیچ یک از شرایط مناسب فوق برقرار نمی باشد، بنابراین برای تقویت مرحله تشخیص اولیه و همچنین حذف غیرخودرو و بهبود خطاهای موجود در خروجی نهایی سیستم، نیازمند یک الگوریتم با یادگیری می باشیم. در بخش بعدی در مورد روش های مبتنی بر یادگیری بحث خواهد شد.

۴-۳ - تقویت تشخیص وسایل نقلیه بر پایه روش های یادگیری

همان‌طور که در بخش‌های قبل ذکر شد، سیستم ردیابی در تقاطع‌ها نیاز به تقویت دارد، تا مشکلات ناشی از پیوستگی خودروها در تصویر خروجی مدل ترکیبی گوسی، وجود غیرخودرو، تشخیص ناقص وسایل نقلیه و... را رفع نماید.

بدین منظور از سیستم ماشین بردار پشتیبان، برای جداسازی خودرو از غیرخودرو استفاده خواهد شد. [۱۹]

ماشین بردار پشتیبان^۱ یک مجموعه از روشهای یادگیری بانظارت^۲ است که برای طبقه بندی^۳ استفاده می‌شود و طبقه بندی را با استفاده از ابرصفحه‌های خطی برای تصمیم‌گیری در فضای ویژگی انجام می‌دهد. در طول آموزش، ابرصفحه‌ها بمنظور تفکیک داده‌های آموزشی با برجسب‌های مختلف محاسبه می‌شوند. اگر داده‌های آموزشی بصورت خطی قابل تفکیک نباشند، می‌توان از یک تابع کرنل برای تبدیل داده‌ها به یک فضای بردار جدید استفاده نمود. داده‌ها می‌باید در فضای جدید بصورت خطی قابل تفکیک شوند. ماشین بردار پشتیبان عملکرد بسیار خوبی برای مجموعه داده‌های آموزشی بزرگ دارد و در این روش پیچیدگی آموزش متناسب با تعداد نمونه‌های آموزشی افزایش می‌یابد.

ورودی SVM، بردارهای ویژگی می‌باشد. از آنجاییکه در تقاطع‌ها، خودروها از جهات مختلفی وارد می‌شوند، و همچنین در ویدئوهای مختلف، ابعاد خودروها با توجه به محل قرارگیری دوربین‌های ترافیکی، کوچک و بزرگ می‌باشد، نیازمند یک ویژگی بسیار قوی برای توصیف خودرو در شرایط فوق می‌باشیم. برای این منظور از توصیف‌گر هیستوگرام گرادیان‌های جهت‌دار (HOG) استفاده خواهد شد. [۵۱]

همچنین یکی از مشکلاتی که در این تحقیق مورد بررسی قرار خواهد گرفت، همپوشانی تصویری خودروها بوده که فرآیند ردیابی را با مشکل مواجه می‌سازد. در این تحقیق با استفاده از یک سیستم

¹ Support vector machine(SVM)

² Supervised learning

³ classification

ساده، دو خودرو را که با هم همپوشانی جزئی داشته‌اند، تشخیص می‌دهیم. سیستم استفاده شده برای حالاتی که بیش از دو خودرو با هم همپوشانی دارند، و یا حالاتی که همپوشانی کامل اتفاق بیافتد، جوابگو نمی‌باشد. در ادامه به بررسی سیستم کلاسه‌بند و تشخیص همپوشانی، برای تقویت مرحله تشخیص، می‌پردازیم.

۴-۳-۱ کلاسه‌بند SVM با استفاده از توصیف‌گر HOG

همانطور که در بخش‌های قبلی توضیح داده شده است، قصد داریم با استفاده از کلاسه‌بند SVM مدلی را برای تقویت سیستم تشخیص وسایل نقلیه متحرک، ارائه نماییم. سیستم‌های مبتنی بر یادگیری شامل دو بخش آموزش و تست می‌باشند. بدین صورت که در ابتدا سیستم را با استفاده از داده‌های آموزشی، آموزش می‌دهیم و سپس از همان مدل آموزشی برای تست، استفاده خواهیم کرد.

همان طوری که در بخش قبلی توضیح داده شد، ما در این تحقیق از توصیف‌گر HOG برای ورودی کلاسه‌بند SVM استفاده خواهیم کرد. هر چه تعداد داده‌های آموزشی بیشتر باشد، سیستم آموزش بهتری دیده و در نتیجه پاسخ مناسب تری را در مرحله تست خواهد داشت. پس از اینکه سیستم SVM را آموزش دادیم، خروجی مرحله حذف پس‌زمینه، با استفاده از مدل آموزشی SVM، تست شده و در صورت تایید به بخش ردیابی فرستاده می‌شود. در این بین برخی از نواقص و اشکالات مربوط به خروجی حذف پس‌زمینه، رفع شده و تشخیص مناسب تری را برای ورود به بخش ردیابی خواهیم داشت. در ادامه نحوه انتخاب مدل آموزشی SVM، مورد بررسی قرار داده خواهد شد.

در این تحقیق داده‌های آموزشی از خروجی مرحله حذف پس‌زمینه به صورت خودکار تهیه می‌گردد. تصاویر با اندازه‌های متغیر و ثابت تهیه شده و سپس به دو بخش کلی داده‌های مثبت و منفی تقسیم می‌شود. داده‌های مثبت شامل خودروهای مختلف در ویدئوهای گوناگون است که با درجه تفکیک‌های مختلف، تهیه شده است. در این تصاویر خودروها از زوایای مختلف دیده می‌شوند، همچنین ابعاد خودروها با توجه به موقعیت دوربین، متفاوت می‌باشد. در کلاس تصاویر منفی، عابری

پیاده، دوچرخه و موتورسیکلت وجود دارد و از طرفی داده‌های منفی خاصی بنام منفی تقریبی^۱ که شامل کناره‌های وسایل نقلیه بوده تهیه می‌گردد. دلیل آموزش مدل با این داده‌ها، این است که مدل مجبور خواهد شد تا در برابر حالت‌هایی که خروجی مرحله GMM، تنها قسمتی از خودرو را به صورت ناقص تشخیص می‌دهد، توسط کلاسه‌بند SVM، تصحیح گردد. همچنین دقت سیستم برای تشخیص همپوشانی دو خودرو، بیشتر می‌شود. در ادامه در مورد نحوه جستجو توسط سیستم آموزش داده شده SVM، برای تعیین محل دقیق خودروها توضیحاتی ارائه خواهد شد، و تاثیر استفاده از داده‌های منفی تقریبی، بصورت واضح‌تری بیان خواهد شد. تصاویر نمونه برای بخش آموزشی در زیر نشان داده شده است. همچنین در بخش نتایج آزمایشی، در مورد پایگاه داده‌ی استفاده شده در این تحقیق بحث خواهد شد.



شکل ۴-۶: نمونه تصاویر آموزشی SVM (الف) مثبت (ب) منفی (پ) منفی تقریبی

در روش هیستوگرام گرادیان جهت‌دار، برای محاسبه بردار ویژگی، گرادیان تصویر ورودی به شبکه‌ای از سلول‌ها تقسیم بندی می‌شود. برای هر سلول، یک هیستوگرام از جهت گرادیان یا جهت های لبه‌ی پیکسلها محاسبه می‌شود، طریقه محاسبه هیستوگرام به این صورت می‌باشد که با توجه به پارامتر فاکتور جهت^۲، مقادیر گرادیانی هر پیکسل که بین زوایای ۰ تا ۱۸۰ درجه هستند به چندین بخش مساوی تقسیم‌بندی می‌شود. برای مثال با در نظر گرفتن فاکتور جهتی برابر با ۹ جهت های گرادیانی ۰ تا ۱۸۰ درجه، به ۹ بخش ۲۰ درجه‌ای تقسیم می‌شود. هیستوگرام تشکیل شده بر اساس

¹ Hard Negative

² Orientation bin

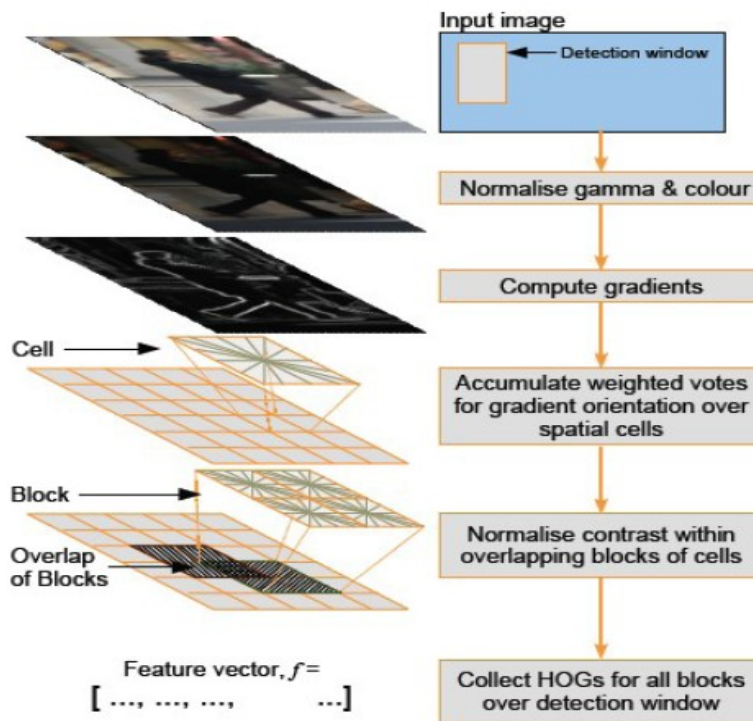
فراوانی مقادیر گرادیانی (جهت و اندازه) در هر بخش زاویه‌ای برای کل سلول محاسبه می‌گردد. این روش کارایی و عملکرد بسیار خوبی در مقایسه با سایر مجموعه ویژگی‌های موجود مانند موجک دارد. فرضیه اساسی در این روش این است که ظاهر و شکل شیء اغلب می‌تواند توسط توزیع گرادیان شدت محلی یا جهت لبه حتی بدون داشتن آگاهی دقیق از گرادیان متناظر یا مکان لبه، نسبتاً بهتر تشخیص داده شود. در مقاله اصلی، این بردار برای شناسایی عابران پیاده استفاده شده است. در ادامه بحث این مفهوم برای شناسایی وسایل نقلیه استفاده می‌شود.

به منظور تغییر ناپذیری از روشنایی و سایه، نرمالیزه کردن کنتراست مفید است. از این رو هیستوگرام محلی را در نواحی مکانی بزرگتری به نام بلوک‌ها محاسبه و از نتایج آن‌ها برای نرمالیزه کردن همه سلول‌های آن بلوک استفاده می‌کنیم. و در انتها بردار ویژگی، از گردآوری هیستوگرام‌های گرادیان‌های جهت‌دار مربوط به تمامی بلوک‌ها که در مرحله قبل محاسبه شده است بدست می‌آید.

[۵۱]

همانطور که در بالا ذکر شده است دو پارامتر مهم در توصیف گر HOG، اندازه سلول^۱ و اندازه فاکتور جهت^۲ می‌باشد. در شکل زیر روال کلی بدست آوردن HOG از یک تصویر نشان داده شده است.

^۱ Cell size^۲ Bin size



شکل ۴-۷: خلاصه‌ی روش HOG [۵۲]

با توجه به اندازه‌ی تصویر ورودی به توصیفگر HOG، سایز بردار خروجی متفاوت خواهد بود. در صورتیکه اندازه سلول، برابر یک-چهارم اندازه تصویر در نظر گرفته شود، اندازه بردار خروجی HOG به اندازه تصویر ورودی وابسته نمی‌باشد. به این صورت که با استفاده از یک بلوک 2×2 هر تصویر ورودی با هر سایزی را به چهار بخش تقسیم نماییم. [۵۳]

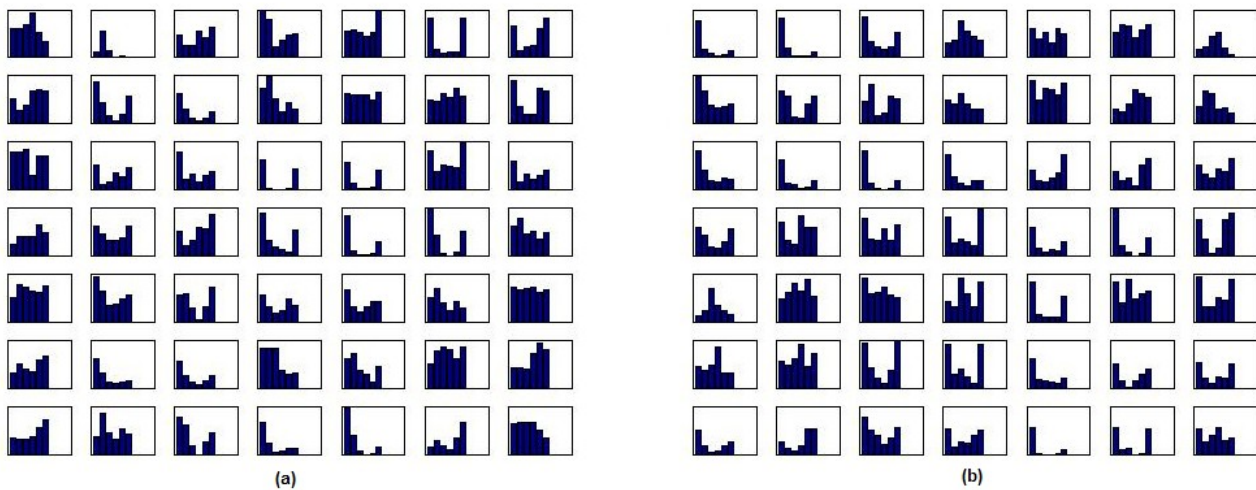
در ابتدا با توجه به اینکه در ویدئوهای استفاده شده، اندازه‌ی خودروها متفاوت می‌باشند، از تصاویر آموزشی با اندازه‌های متفاوت، برای آموزش کلاس‌بند به منظور داشتن قدرت عمل بالاتر استفاده شد. با علم به این مساله که اندازه سلول‌ها در توصیفگر HOG، با توجه به مباحث فوق، یک-چهارم اندازه تصویر آموزشی در نظر گرفته شده بود. اما اندازه‌های در نظر گرفته شده برای هر سلول نسبتاً بزرگ بوده و در نتیجه طول بردارهای خروجی توصیفگر کوچک بوده، و دقت تشخیص مدل SVM، خیلی بالا نبوده، بنابراین اگرچه توانایی در نظر گرفتن اندازه‌های متفاوت داده‌های آموزشی برای کاربرد این تحقیق بسیار مناسب بوده، اما به دلیل عدم دقت خیلی بالا، اندازه‌های ثابتی را برای داده‌های ورودی

در نظر گرفتیم، تا بتوانیم طول بزرگتری را برای بردار توصیف‌گر داشته باشیم، که به همین ترتیب دقت تشخیص مدل SVM، بالاتر باشد. بنابراین در این بخش از توصیف‌گر HOG، با استفاده از داده‌ی ورودی ثابت مدلی با دقت بالا برای کلاسه‌بند خود تعیین می‌نماییم. در بخش تشخیص همپوشانی نیز از همین توصیف‌گر HOG و کلاسه‌بند SVM استفاده خواهد شد. در آن قسمت پس از تعیین محل پنجره همپوشانی، از آنجاییکه اطمینان از وجود خودرو در داخل پنجره همپوشانی داریم، دقت خیلی بالا حائز اهمیت نمی‌باشد، اما استفاده از توانایی جستجو در داخل پنجره همپوشانی با استفاده از اندازه‌های مختلف داده‌ی توصیف‌گر، بسیار مناسب می‌باشد. برای جمع‌بندی مباحث مطرح شده، در این بخش مدل ۱ را برای کلاسه‌بند SVM با استفاده از داده‌های اندازه ثابت، و برای بخش همپوشانی مدل ۲ را برای کلاسه‌بند SVM با استفاده از داده‌های اندازه متغیر خواهیم داشت. نحوه جستجو با استفاده از پنجره‌های لغزان^۱، در داخل بلاب‌های خروجی از مرحله حذف پس‌زمینه خواهد بود، که جزئیات آن در بخش‌های بعدی مطرح می‌گردد.

در این بخش همان‌طور که مطرح شد، پارامترهای مربوط به توصیف‌گر HOG از این قرار در نظر گرفته می‌شود: داده‌ها با اندازه 64×64 با توجه به ویدئوهای استفاده شده، در نظر گرفته شده‌اند، به‌طوری‌که کل وسیله نقلیه را پوشش دهند و از طرفی خیلی بزرگ در نظر گرفته نشده باشند. برای آموزش مدل شماره ۱ SVM، که از داده‌های ثابت ورودی استفاده می‌شود، اندازه‌ی مربوط به سائز سلول‌ها ۸ در نظر گرفته می‌شود. (این اندازه باید بر اندازه مربوط به داده‌ی ورودی بخش‌پذیر باشد [۵۱]) در نظر گرفتن این اندازه برای هر سلول، همان‌طوری‌که در ادامه خواهید دید، دقت بالایی را برای SVM، به همراه خواهد داشت. هر گروه 2×2 از سلول‌ها در داخل یک بلوک قرار می‌گیرند و بلوک‌ها با یکدیگر همپوشانی خواهند داشت. [۵۱].

¹ Sliding Window

مقدار اندازه فاکتور جهت برای دستیابی به بالاترین دقت در ادامه تعیین خواهد شد. در تصویر زیر خروجی قسمت اختصاص هیستوگرام به هر سلول، برای دو نمونه تصویر ورودی نشان داده شده است.



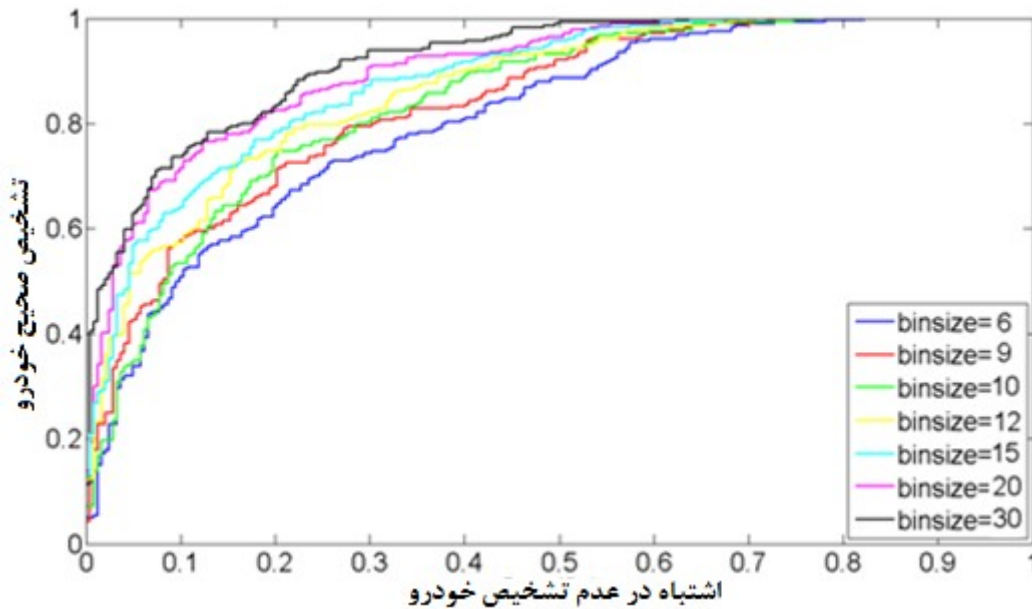
شکل ۴-۸: اختصاص هیستوگرام به هر سلول برای دو نمونه تصویر ورودی

پس از اینکه با استفاده از داده‌های آموزشی مختلف، مقادیر بردارهای توصیف‌گر HOG بدست آمد. از آنها برای آموزش مدل شماره ۱ کلاسه‌بند SVM، استفاده می‌نماییم.

در ابتدا تصاویر زیادی به صورت خودکار از ویدئوهای مختلف با سایز 64×64 تهیه گردید. اما از آنجاییکه استفاده از اندازه سلول برابر ۸، بردارهای بزرگی را برای آموزش SVM، تولید می‌کرد، با در نظر گرفتن تعداد بالای داده‌ی آموزشی، ماتریس خروجی توصیف‌گر HOG حجم بالایی پیدا کرده و کتابخانه SVM نرم افزار متلب قادر به کار کردن با آن حجم بالا نبود، به همین دلیل تنها قادر به آموزش مدل ۱ کلاسه‌بند با ۲۷۰ داده‌ی آموزشی بوده‌ایم. در ادامه مراحل آموزش کلاسه‌بند توضیح داده می‌شود.

مدل شماره ۱ با استفاده از ۱۷۰ تصویر آموزشی منفی و ۱۰۰ داده‌ی مثبت، آموزش داده شده و سپس با استفاده از همین تعداد داده، با اعمال اندازه‌های مختلفی از فاکتور جهت، ارزیابی شد.

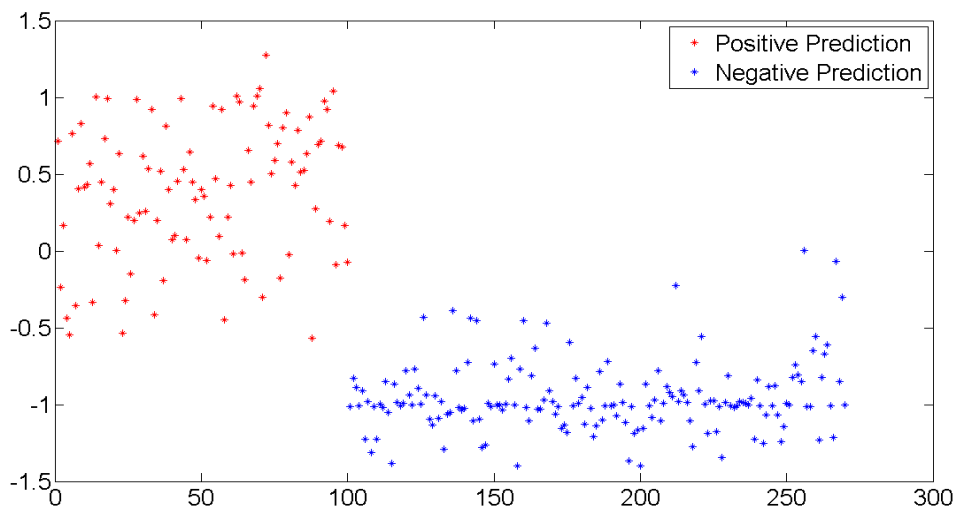
در تصویر زیر نمودار ROC مربوط به ارزیابی مدل ۱ کلاسه‌بند بر اساس نسبت (TP/FN) ^۱ با توجه به مقادیر مختلف اندازه فاکتور جهت نمایش داده شده است.



شکل ۴-۹: منحنی ROC مربوط به مدل آموزشی شماره ۱ با توجه به اندازه‌های مختلف فاکتور جهت همانطور که در شکل فوق دیده می‌شود، هرچه مقدار اندازه فاکتور جهت بیشتر می‌شود، منحنی ROC، خروجی مناسب‌تری دارد.

با توجه به هر داده‌ی ارزیابی، مقادیر مختلفی برای احتمال مدل کلاسه‌بند SVM، محاسبه می‌شود. در شکل زیر مقادیر این احتمال برای دو کلاس مثبت و منفی برای حالتی که از مقدار فاکتور جهت ۶ استفاده شده است، نشان داده شده است. سایر پارامترها مانند حالت اولیه تنظیم شده‌اند، مقدار اندازه سلول برابر ۸ در نظر گرفته شده است و هر گروه ۲۴ سلول‌ها در یک بلوک با همپوشانی قرار می‌گیرند.

^۱ True Positive / False Negative



شکل ۴-۱۰: نمایش خروجی احتمالات مدل SVM برای تعیین میزان آستانه بین دو کلاس مثبت و منفی

همانطور که در شکل فوق دیده می‌شود، به دلیل وجود داده‌های منفی تقریبی، در برخی از احتمالات، مدل SVM معرفی شده مقادیر کلاس منفی و مثبت را نزدیک به هم تخمین زده است، با توجه به شکل فوق ما مقدار آستانه $0/2$ را برای جداسازی بین کلاس‌های خودرو و غیر خودرو در نظر می‌گیریم. در ادامه جدول مربوط به ارزیابی مدل SVM شماره ۱ آورده شده است:

جدول ۴-۱: ارزیابی مدل SVM شماره ۱

Dataset	Train Pos	Train Neg	Test Pos	Test Neg	Cell size	Bin size	Feature vector size	ROC Area	Accuracy
2	100	170	170	100	8	6	864	81.3	0.9148
2	100	170	170	100	8	9	1296	83.87	0.9037
2	100	170	170	100	8	12	1728	84.43	0.8852
2	100	170	170	100	8	15	2160	86.53	0.8667
2	100	170	170	100	8	18	2592	88.21	0.837
2	100	170	170	100	8	20	2880	90.18	0.837
2	100	170	170	100	8	30	4320	92.76	0.7333

همانطور که در جدول ۴-۱ دیده می‌شود با افزایش مقدار اندازه فاکتور جهت، اندازه‌ی بردار ویژگی افزایش یافته، و منحنی ROC، پاسخ مناسب‌تری خواهد داشت، اما دقت سیستم در حالتی که اندازه فاکتور جهت ۶ در نظر گرفته شود، بالای ۹۱/۴۸٪ خواهد بود. به همین دلیل ما نیز از همین مدل برای تست در ادامه فرآیند تشخیص استفاده خواهیم کرد.

روال کلی سیستم فوق در تقویت مرحله‌ی تشخیص بدین صورت است که در خروجی مرحله حذف پس‌زمینه، بلاب‌هایی که نشانگر محل وسایل نقلیه هستند، نمایش داده می‌شوند. تعداد بلاب‌ها مشخص شده و با استفاده از پنجره لغزان در داخل بلاب‌ها به جستجو پرداخته تا بهترین تطبیق برای محل خودرو پیدا شود. در مواقعی که احتمال مدل کلاسه‌بند از مقدار آستانه کمتر باشد، آن بلاب به قسمت ردیابی فرستاده نخواهد شد. در برخی از موارد اندازه بلاب بسیار کوچک می‌باشد و این حالت بیانگر تشخیص ناقص در مرحله قبل می‌باشد. اگر مقدار احتمال مناسب باشد، اندازه آن بلاب برابر با اندازه پنجره جستجو که در این مدل برابر ۶۴×۶۴ است، برای واحد ردیابی فرستاده می‌شود. در برخی از موارد نیز در ویدئوهای بسیار شلوغ و پر نویز، بلاب‌های مربوط به وسایل نقلیه‌ی نزدیک به هم، به یکدیگر پیوسته می‌شوند، که در این شرایط به ازای دو محلی که پنجره‌های لغزان داخل آن بلاب، بیشترین احتمال را داشته باشند، آن بلاب به دو بخش تقسیم شده و به عنوان دو وسیله نقلیه برای واحد ردیابی فرستاده می‌شود، البته شایان ذکر است که دو پنجره لغزان نمی‌بایست هیچ تداخلی با یکدیگر داشته باشند. بدین ترتیب استفاده از مدل شماره ۱ کلاسه‌بند پیشنهادی، در رفع مشکلاتی مانند حذف غیرخودرو، رفع تشخیص ناقص، رفع بهم پیوستگی بلاب‌ها کمک‌کننده می‌باشد. در بخش نتایج آزمایشی علاوه بر ارائه‌ی مقادیر پارامترها و معرفی پایگاه داده، نتایج مربوط به این مدل مورد بررسی قرار می‌گیرد.

۲-۳-۴ تشخیص همپوشانی

در تحلیل ترافیک در تقاطع‌ها با مشکلاتی از قبیل همپوشانی جزئی و کلی مواجه هستیم، که باعث به وجود آمدن خطا در سیستم‌های تحلیل ترافیک می‌گردد. در این بخش قصد داریم با استفاده از سیستم تشخیص همپوشانی، همپوشانی جزئی موجود بین دو وسیله نقلیه متحرک را تشخیص دهیم. و پس از تعیین محل دقیق هر کدام از دو وسیله نقلیه، آن موقعیت را برای سیستم ردیابی ارسال نماییم. سیستم پیشنهادی از شکل بلاب‌ها (نواحی تشخیص داده شده در خروجی GMM)، و مشخصات مورفولوژیکی آن‌ها مانند استواری^۱، جهت تشخیص همپوشانی استفاده خواهد نمود. روش‌های مبتنی بر رنگ نیز می‌توانند، با توجه به تفاوت سطوح رنگی، جهت تشخیص همپوشانی استفاده شود، اما در شرایطی که دو خودروی دارای همپوشانی، رنگ‌های یکسانی داشته باشند، این روش کارایی چندانی ندارد.

در کل خودروها شکل محدب^۲ گونه‌ای دارند. به همین دلیل فرض می‌شود، شکل بلاب دو خودرویی که با هم همپوشانی دارند، تحدب کمتری خواهد بود. همچنین این امکان وجود دارد که دو خودرو با یکدیگر به شدت همپوشانی داشته باشند و شکل بلاب مشترک آنها باز هم محدب دیده شود، در چنین مواقعی، سیستم فوق نمی‌تواند همپوشانی کامل را تشخیص دهد، چراکه اطلاعات بسیار کمی در مورد خودروهایی که با هم همپوشانی کامل دارند، موجود می‌باشد. [۵۴]

استواری به عنوان مساحت یک بلاب که بروی مساحت فضای محدب^۳ مربوط به آن بلاب تقسیم می‌گردد تعریف می‌شود. فرض می‌شود که در کل N بلاب وجود دارد، B_i به بلاب شماره i اشاره دارد و C_i فضای محدب داخلی آن بلاب می‌باشد. در این صورت استواری بلاب شماره i با S_i نشان داده خواهد شد و به صورت زیر تعریف می‌شود:

¹ Solidity

² Convex

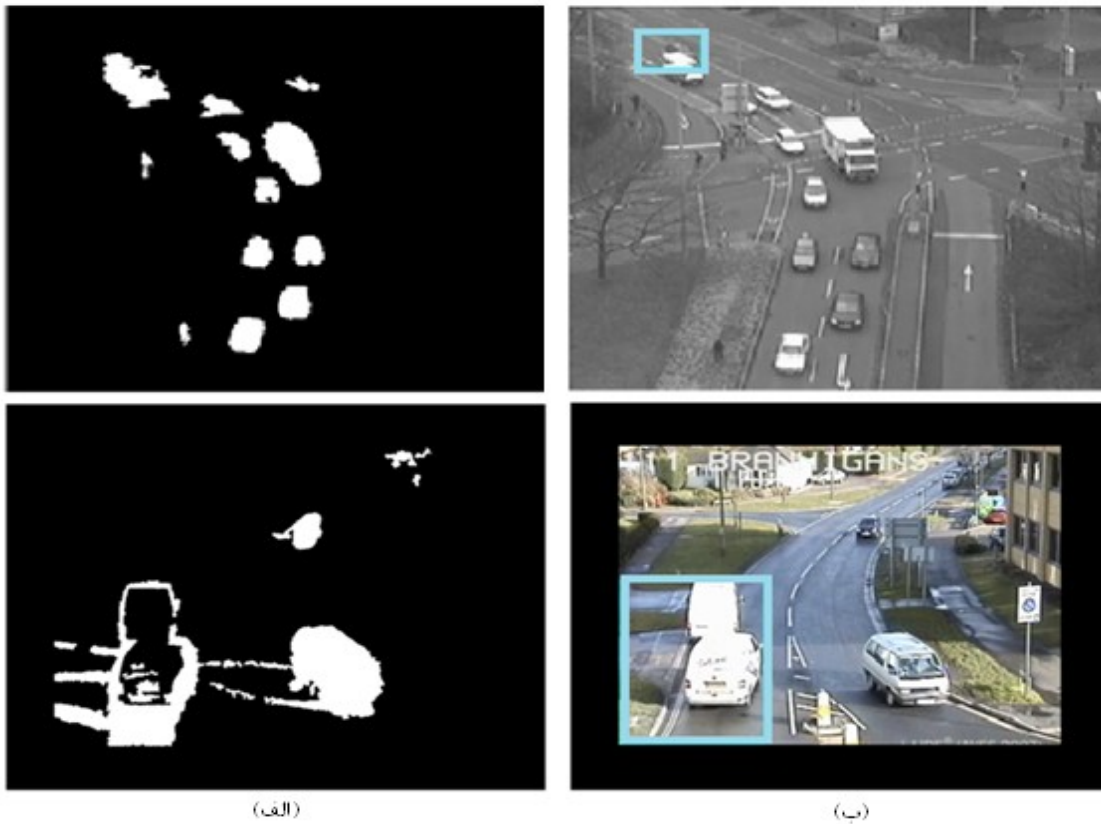
³ Convex hull

$$S_i = \text{Area}(B_i) / \text{Area}(C_i) \quad (۵-۴)$$

در صورتیکه $(S_i < T_s)$ باشد، فرض می‌کنیم بلاک شماره i شامل بیش از یک وسیله نقلیه است.

در این رابطه T_s آستانه مربوط به میزان استواری است، که به صورت تجربی تعیین می‌گردد.

در تصاویر زیر بلاک‌های شامل دو وسیله نقلیه با استفاده از سیستم فوق تشخیص داده شده‌اند.

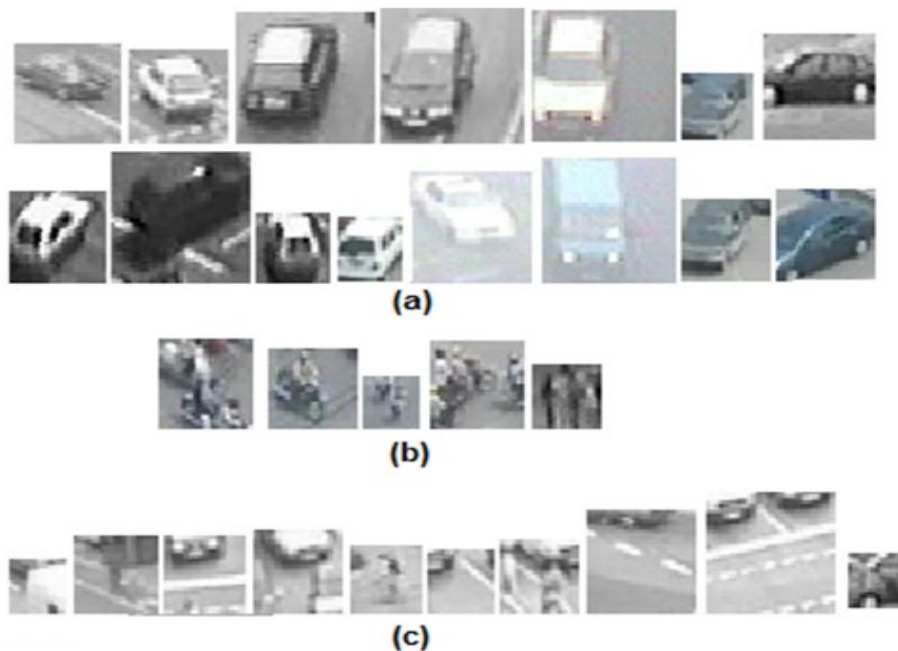


شکل ۴-۱۱: (الف) تصاویر خروجی مرحله GMM (ب) تشخیص همپوشانی

در بخش نتایج آزمایشی، تاثیر بخش تشخیص همپوشانی در تصحیح عملکرد سیستم ردیابی، مورد

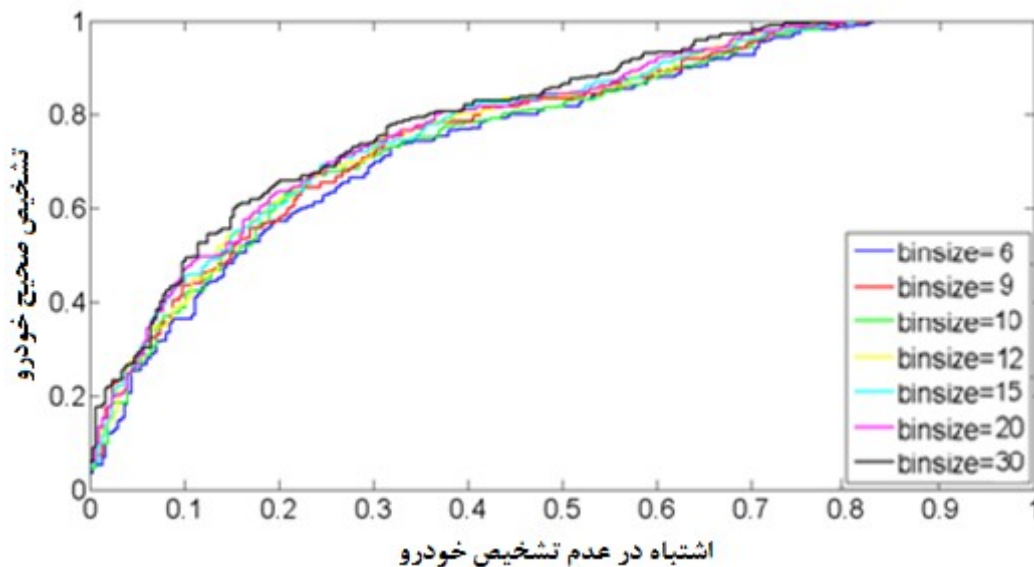
بررسی قرار خواهد گرفت.

در ادامه قصد داریم، در مورد جداسازی خودروها در داخل بلاب همپوشانی بحث نماییم. همانطور که در بخش های قبلی آورده شده است، برای این کار از سیستم HOG/SVM با مدل آموزشی شماره ۲، استفاده می شود. در این مدل داده های ورودی اندازه های متفاوتی داشته، تا توانایی سیستم برای جستجو در داخل بلاب مشکوک به همپوشانی بیشتر شود. و بتوانیم بهترین تطبیق را با توجه به اندازه خودرو، داشته باشیم که در حالتی که داده های اندازه ثابت استفاده شود، سیستم این توانایی را ندارد. همانطور که در بخش های قبل آورده شده است، در این حالت برای اینکه اندازه بردار ویژگی HOG، برای داده های با اندازه های مختلف، یکسان باشد، ناگزیر هستیم که اندازه سلول که از پارامترهای اصلی توصیف گر HOG است را برابر یک-چهارم اندازه تصویر در نظر بگیریم. [۵۳] در این حالت نیز، همانند مدل آموزشی شماره ۱، هر چهار سلول با هم به صورت ۲x۲، در یک بلوک با همپوشانی قرار می گیرند. در شکل زیر تعدادی از تصاویر نمونه آموزشی مثبت، منفی و منفی تقریبی، که دارای اندازه های متفاوتی می باشند، نمایش داده شده است:



شکل ۴-۱۲: تصاویر نمونه آموزشی با اندازه های متفاوت (a) مثبت (b) منفی (c) منفی تقریبی

مراحلی که در بخش قبل برای بدست آوردن مدل کلاسه‌بند انجام شد، در این مرحله نیز تکرار می‌گردد و سپس با توجه به تغییرات فاکتور جهت، در این مرحله نیز منحنی ROC، بدست می‌آید. در مرحله قبلی از آنجاکه اندازه سلول برابر با ۸ بود و طول بردار HOG بسیار بزرگ می‌شد، متلب قادر به تحلیل روی داده‌های فراوان بدست آمده نبود. در این مرحله چنین مشکلی وجود نداشته و می‌توان هر تعداد از داده‌های ورودی را برای آموزش مدل شماره ۲ کلاسه‌بند استفاده نمود. منحنی ROC در شکل زیر نمایش داده شده است:



شکل ۴-۱۳: منحنی ROC مربوط به مدل آموزشی شماره ۲ با توجه به اندازه‌های مختلف فاکتور جهت

در کل، اگرچه تعداد داده‌های بیشتری برای آموزش این مدل استفاده شده است، ولی مساحت زیر منحنی ROC نسبت به مدل قبلی کمتر می‌باشد. دلیل این مساله به طول بالای بردار ویژگی برمی‌گردد، که تاثیر زیادی در بالا رفتن عملکرد مدل دارد.

در منحنی ROC فوق، مانند حالت قبلی با افزایش مقدار فاکتور جهت، منحنی ROC عملکرد مناسب‌تری را از خود نشان می‌دهد. در ادامه جدول مربوط به عملکرد مدل شماره ۲ کلاسه‌بند SVM که برای جداسازی خودروها در داخل بلاک همپوشانی، استفاده می‌گردد، نشان داده خواهد شد.

جدول ۴-۲: ارزیابی مدل SVM شماره ۲

Dataset	Train Pos	Train Neg	Test Pos	Test Neg	Cell size	Bin size	Feature vector size	ROC Area	Accuracy
1	600	400	300	312	Mask size/4	6	96	74.27	66.62%
1	600	400	300	312	Mask size/4	9	144	75.45	55.23%
1	600	400	300	312	Mask size/4	12	192	76.17	59.67%
1	600	400	300	312	Mask size/4	15	240	76.74	52.87%
1	600	400	300	312	Mask size/4	18	288	77.07	55.44%
1	600	400	300	312	Mask size/4	20	320	77.51	53.47%
1	600	400	300	312	Mask size/4	30	480	78.72	54.23%

همانطور که در جدول فوق دیده می‌شود، تعداد ۶۰۰ عدد داده‌ی مثبت و ۴۰۰ عدد داده‌ی منفی برای آموزش مدل شماره ۲، استفاده شده است. این تصاویر به صورت اتوماتیک از خروجی مرحله تشخیص اولیه، بدست آمده‌اند و سپس به صورت با سرپرست، به دو بخش مثبت و منفی تقسیم شده‌اند. اندازه این داده‌ها متغیر می‌باشد.

تعداد داده‌های استفاده شده برای ارزیابی سیستم، برابر با ۳۰۰ داده‌ی مثبت و ۳۱۲ عدد داده‌ی منفی می‌باشد. دقت سیستم، به ازای اندازه‌های متفاوتی که برای فاکتور جهت در نظر گرفته شده است، بدست می‌آید. بالاترین دقت مانند حالت قبلی برای فاکتور جهت برابر با ۶ بدست آمده است، از این رو ما از همین مدل برای جداسازی خودروها در بلاب همپوشانی استفاده خواهیم نمود.

تا اینجا کار با توجه به شکل ۴-۱ که شمای کلی سیستم را نشان می‌دهد، عملیات تشخیص اولیه وسایل نقلیه متحرک و عابرین پیاده و موتورسیکلت و دوچرخه با استفاده از الگوریتم حذف پس

زمینه با روش GMM، و سپس حذف سایه و نویز از روی آن انجام شد. در ادامه برای تقویت مرحله قبلی از سیستم تشخیص همپوشانی و سیستم HOG/SVM استفاده شد، تا کلیه مشکلات مرحله قبلی را از بین برده و خروجی مناسبی را برای مرحله ردیابی داشته باشد. در مرحله ردیابی با استفاده از فیلتر کالمن و الگوریتم MHT، وسایل نقلیه متحرک در هر فریم زمانی دنبال می‌شوند. با توجه به حرکت تقریباً خطی و تعداد زیاد خودروها در تقاطع‌ها، الگوریتم ردیابی Kalman/MHT به‌عنوان بهترین روش در این تحقیق مورد استفاده قرار خواهد گرفت. فیلتر های بسیاری برای ردیابی وسایل نقلیه وجود دارد که یکی از قوی ترین آنها، فیلتر ذره‌ای می‌باشد. فیلتر ذره‌ای در ردیابی اجسامی که پیچیدگی حرکتی بالایی دارند، بسیار موثر عمل می‌کند، اما از سوی دیگر با افزایش تعداد اجسام متحرک، فیلتر ذره‌ای بسیار کند عمل کرده و نمی‌توان از آن برای یک سیستم بلادرنگ استفاده نمود. در صورتیکه فیلتر کالمن در تخمین و ردیابی حرکتی اجسامی که حرکتی خطی دارند، بهینه‌ترین روش می‌باشد. با توجه به مباحث فوق ما از سیستم ردیابی کالمن به همراه MHT، برای ردیابی همزمان تعداد بالایی از وسایل نقلیه متحرک استفاده نمودیم. در بخش بعدی در مورد پایگاه داده استفاده شده، پارامترهای مختلف هر زیر سیستم بحث شده و در نهایت با نمایش عملکرد سیستم در هر بخش، جدول نهایی ارزیابی سیستم را مورد بررسی قرار می‌دهیم.

۴-۴ - نتایج آزمایشی

در این بخش به بررسی نتایج سیستم پیشنهادی می‌پردازیم. در مورد پایگاه داده‌ی استفاده شده در این تحقیق بحث شده، مقادیر پارامترهای مختلف را بررسی کرده و همچنین کتابخانه‌ها و نرم‌افزارهای استفاده شده در پیاده‌سازی‌ها مورد بررسی قرار می‌گیرند.

پایگاه داده‌ی اصلی استفاده‌شده در این تحقیق از داده‌های مربوط به تقاطع‌ها که توسط دانشگاه Karlsruhe [۵۵] تهیه شده است استفاده می‌نماید. ویدئوهای قرار داده شده در این پایگاه داده شامل ۱۲ ویدئوی گوناگون از تقاطع‌های مختلف می‌باشد. چندین ویدئو از یک تقاطع در زمان‌های مختلف

تهیه شده است. همچنین در فصل‌های مختلف و شرایط جوی مختلف نیز ویدئوهایی قرار داده شده است، که ما به‌غیر از شرایط معمولی آب و هوایی، از ویدئوهای با وجود مه شدید هم استفاده نمودیم. خروجی سیستم پیشنهادی عملکرد بالایی در تشخیص و ردیابی در شرایط جوی با مه غلیظ، از خود نشان می‌دهد. در بخش‌های آتی به نمایش خروجی سیستم پیشنهادی با استفاده از ویدئوهای مختلف می‌پردازیم.

ویدئوهای آورده شده در پایگاه داده‌ی دانشگاه Karlsruhe، با درجه تفکیک‌های پایین می‌باشند که عملیات تشخیص و ردیابی را مشکل می‌سازد، پایین‌ترین درجه تفکیک مربوط به یک ویدئوی سیاه و سفید با کیفیت 191×256 می‌باشد و بالاترین کیفیت ویدئوی استفاده شده از این پایگاه داده، مربوط به یک ویدئوی رنگی با درجه تفکیک 576×768 به همراه مه غلیظ می‌باشد. نمونه‌ای از این تصاویر در شکل زیر دیده می‌شود:



شکل ۴-۱۴: نمونه تصاویری از ویدئوهای پایگاه داده استفاده شده در این پایان‌نامه

علاوه بر ویدئوهای فوق از یکی از ویدئوهای مربوط به پایگاه داده i-LIDS و مربوط به قسمت AVSS 2007، استفاده شده است. کیفیت این ویدئو ۵۷۶ × ۷۲۰ و رنگی می‌باشد. [۵۶]

همچنین در این تحقیق، از برخی از ویدئوهای ترافیکی که از تقاطع‌های شهری داخل کشور با درجه تفکیک ۵۲۰ × ۶۴۰ تهیه شده است، استفاده شده است.

استفاده از ویدئوهای مختلف با کیفیت‌های گوناگون، در شرایط جوی مختلف، با وجود محل قرارگیری دوربین‌ها در فاصله و جهات مختلف، باعث شده است که در این تحقیق، سیستم پیشنهادی کلی‌ترین حالات ممکن را مورد بررسی قرار دهد. شایان ذکر است که در حالات خاص که تنها شرایط مشخصی بررسی می‌گردد، الگوریتم‌های ارائه شده، می‌توانند با توانایی بیشتری آموزش دیده و دقت بالاتری را برای سیستم نهایی، داشته باشند.

در پیاده‌سازی بخش‌های مختلف این تحقیق، برخی از کتابخانه‌ها و نرم‌افزارهای آماده استفاده شده است، که در ادامه ذکر می‌گردند: در پیاده‌سازی حذف پس‌زمینه با روش GMM و همچنین قسمتی از پیاده‌سازی مربوط به فیلتر کالمن از کتابخانه آماده در [۵۷] استفاده شده است و برای پیاده‌سازی کلاس‌بند SVM از کتابخانه معروف light SVM [۵۸] استفاده شده است. تمامی پیاده‌سازی‌های فوق با استفاده از زبان‌های خانواده C نوشته شده است که باعث شده است، سرعت سیستم پیشنهادی بالا رفته و پیاده‌سازی شبه بلادرنگ را داشته باشیم. سایر پیاده‌سازی‌ها در محیط نرم‌افزار متلب نوشته شده است. و با استفاده از دستور تابع Mex، ارتباط با کتابخانه برنامه‌های C، برقرار شده است. در ادامه این بخش جزئیات مربوط به پیاده‌سازی کل سیستم مورد بررسی قرار خواهد گرفت. پس از اعمال الگوریتم‌های GMM و حذف نویز و سایه، که تشخیص اولیه‌ای را برای ما به دنبال دارد، به دلیل مشکلاتی از قبیل وجود غیر خودرو، پیوستگی برخی از بلاب‌ها به یکدیگر و تشخیص ناقص خودرو، از یک سیستم مبتنی بر یادگیری، برای رفع این مشکلات استفاده شده است.

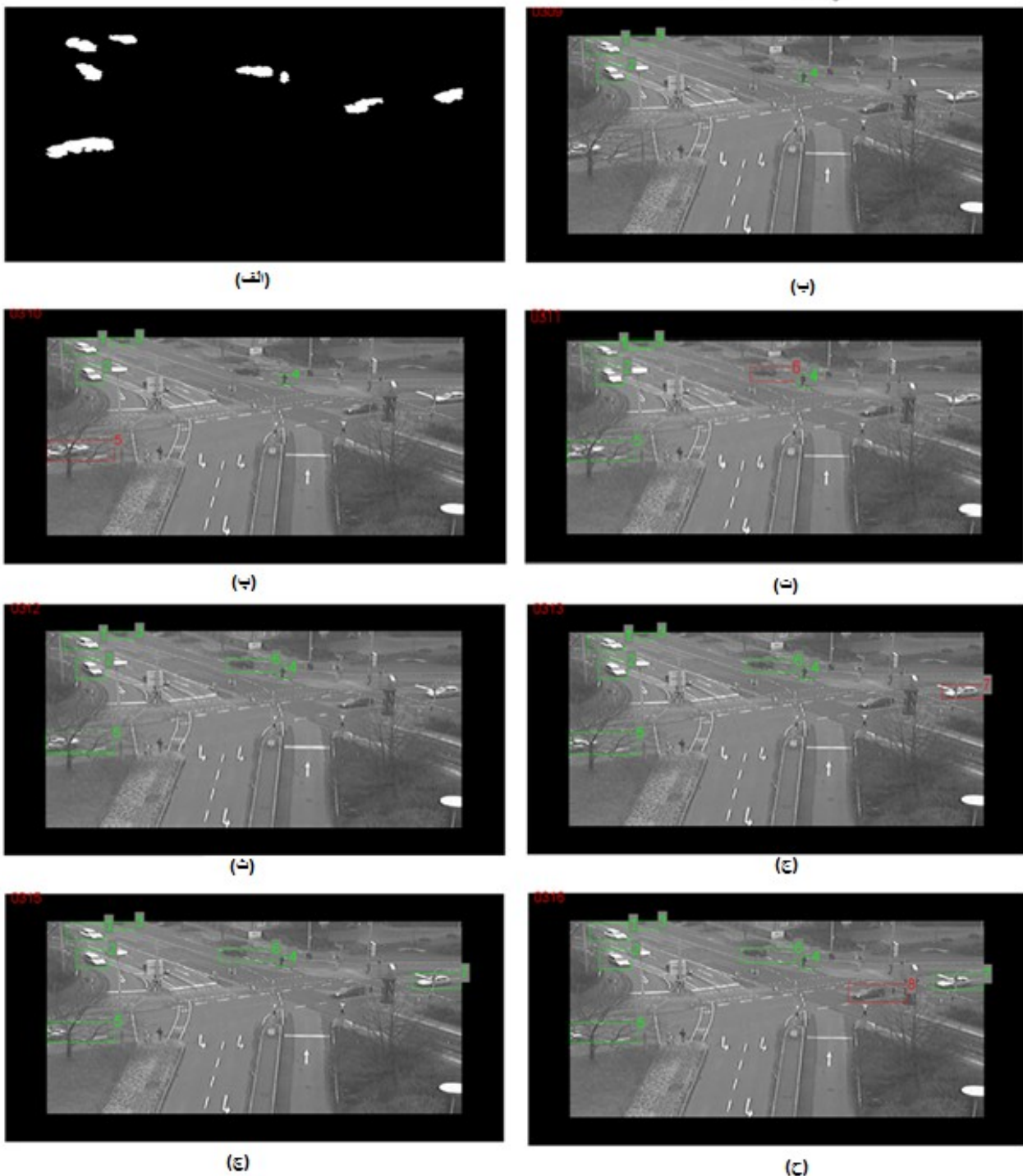
در ابتدا در سیستم تشخیص همپوشانی، از مشخصات مورفولوژیکی بلاب‌ها برای تعیین همپوشانی

استفاده شده است. آستانه تشخیص استواری به طور تجربی برابر با ۰.۶۴ در نظر گرفته شده است. در هر فریم این عملیات برای تمامی بلاب‌ها تکرار می‌شود. در ابتدا بلابی را در نظر می‌گیریم که همپوشانی در آن تشخیص داده نشده باشد. برای بررسی این بلاب از مدل کلاسه‌بند شماره ۱ که با استفاده از تصاویر با اندازه‌های ۶۴×۶۴ ساخته شده است، استفاده می‌گردد. در این مرحله اگر غیر خودرو در بلاب بود، با توجه به آستانه احتمال ۰.۲، که در بخش قبلی بررسی شد، این بلاب حذف شده و برای خروجی فرستاده نمی‌شود. نحوه جستجو در داخل هر بلاب با استفاده از پنجره لغزان با اندازه ثابت انجام شده به طوری که پیکسل مرکزی پنجره لغزان بروی کلیه سطرها و ستونهای آن بلاب حرکت داده شده، و سپس با توجه به بالاترین احتمالات، بهترین محل تطبیق برای وجود خودرو در داخل آن بلاب بررسی می‌گردد.

علاوه بر وجود غیر خودرو، حالاتی از قبیل تشخیص ناقص و تکه‌ای نیز اتفاق می‌افتد. معمولاً در چنین شرایطی، مساحت آن بلاب کوچک بوده، و حداقل پنجره لغزان ۶۴×۶۴ که برابر با اندازه کوچک‌ترین وسایل نقلیه در پایگاه داده‌های کار گرفته شده هستند، برای واحد خروجی فرستاده می‌شود. با این کار، حالاتی را که تشخیص ناقص داریم، بهبود می‌یابد. در شرایطی که مساحت بلاب‌ها کوچک نیست، همان بلاب برای واحد ردیابی فرستاده می‌شود، چراکه شمای وسیله نقلیه در این حالت از بین نرفته و ردیابی بهتری با استفاده از فیلتر کالمن خواهیم داشت.

در برخی از حالات به هم پیوستگی بلاب‌ها را داریم و این در حالی است که سیستم همپوشانی آن را تشخیص نمی‌دهد. برای رفع این مشکل، مکان‌هایی که دو پنجره لغزان بالاترین احتمال را دارند، به شرطی که با هم تداخلی نداشته باشند، به عنوان دو خودرو تشخیص داده می‌شود و به صورت مجزا برای واحد ردیابی فرستاده می‌شود. در ادامه کلیه موارد فوق نمایش داده می‌شود و در نهایت به ارزیابی سیستم نهایی می‌پردازیم.

در ابتدای پیاده سازی‌ها، ردیابی را بدون در نظر گرفتن سیستم مبتنی بر یادگیری، و تنها با استفاده از GMM، نشان می‌دهیم. و در ادامه، خروجی ردیابی را با توجه به سیستم پیشنهادی نمایش داده، و نقاط قوت سیستم پیشنهادی را بررسی می‌نماییم.



شکل ۴-۱۵: ردیابی بدون در نظر گرفتن سیستم مبتنی بر یادگیری (الف) خروجی GMM (ب تا ح) ردیابی در فریم‌های ۳۰۹ تا ۳۱۶

همان‌طور که در خروجی‌های فوق مشاهده می‌گردد، سیستم ردیابی بدون در نظر گرفتن الگوریتم پیشنهادی، نمایش داده شده است. در شکل‌های فوق بلاب شماره ۴ مربوط به یک دوچرخه‌سوار است، که در سیستم فوق ردیابی شده است. در ادامه همین خروجی با استفاده از، سیستم پیشنهادی نشان داده خواهد شد: در تصاویر ذیل کادرهای قرمز رنگ، مربوط به ردیابی اولیه توسط فیلتر کالمن است و کادرهای سبزرنگ مربوط به ردیابی در فریم‌های بعدی می‌باشد. همان‌طور که دیده می‌شود، سیستم پیشنهادی، بلاب مربوط به دوچرخه‌سوار را حذف نموده است. عملیات شمارش نیز در هر فریم زمانی انجام می‌گردد.



(الف)



(ب)



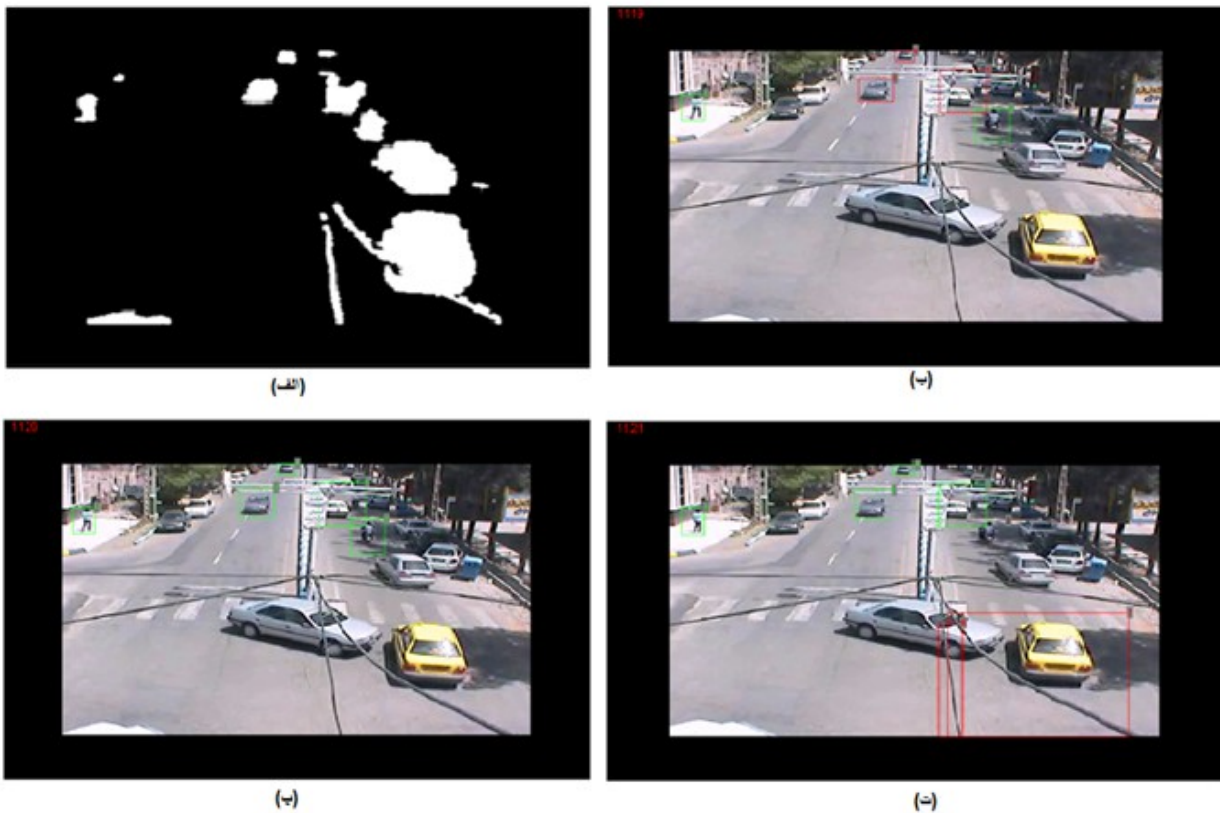
(پ)



(ت)

شکل ۴-۱۶: (الف تا ت) ردیابی در چند فریم با استفاده از سیستم مبتنی بر یادگیری در روش پیشنهادی

در شکل ۴-۱۷ یکی دیگر از حالاتی که سیستم پیشنهادی، غیر خودرو را حذف کرده است، نمایش داده می‌شود، در این ویدئو زاویه دید دوربین و اندازه‌ی خودروها متفاوت می‌باشد، که قدرت الگوریتم پیشنهادی در بررسی کلیه حالات را نشان می‌دهد.



شکل ۴-۱۷: ردیابی بدون در نظر گرفتن سیستم پیشنهادی (الف) خروجی GMM (ب تا ت) ردیابی در فریم‌های متوالی

در ادامه تصاویر مربوط به خروجی پس از استفاده از سیستم پیشنهادی، نمایش داده می‌شود. در کلیه تصاویر نشان داده شده در این بخش، برخی از خودروها، در کلیه‌ی فریم‌ها حرکتی ندارند، بنابراین در خروجی مرحله GMM، قرار نگرفته و ردیابی نخواهند شد. این مورد ممکن است برای خواننده ابهام ایجاد نماید، چراکه از جزئیات حرکتی در هر فریم، مطلع نیست، به همین علت در ابتدای پیاده‌سازی‌ها معمولاً یک خروجی از مرحله GMM، نشان داده شده است.

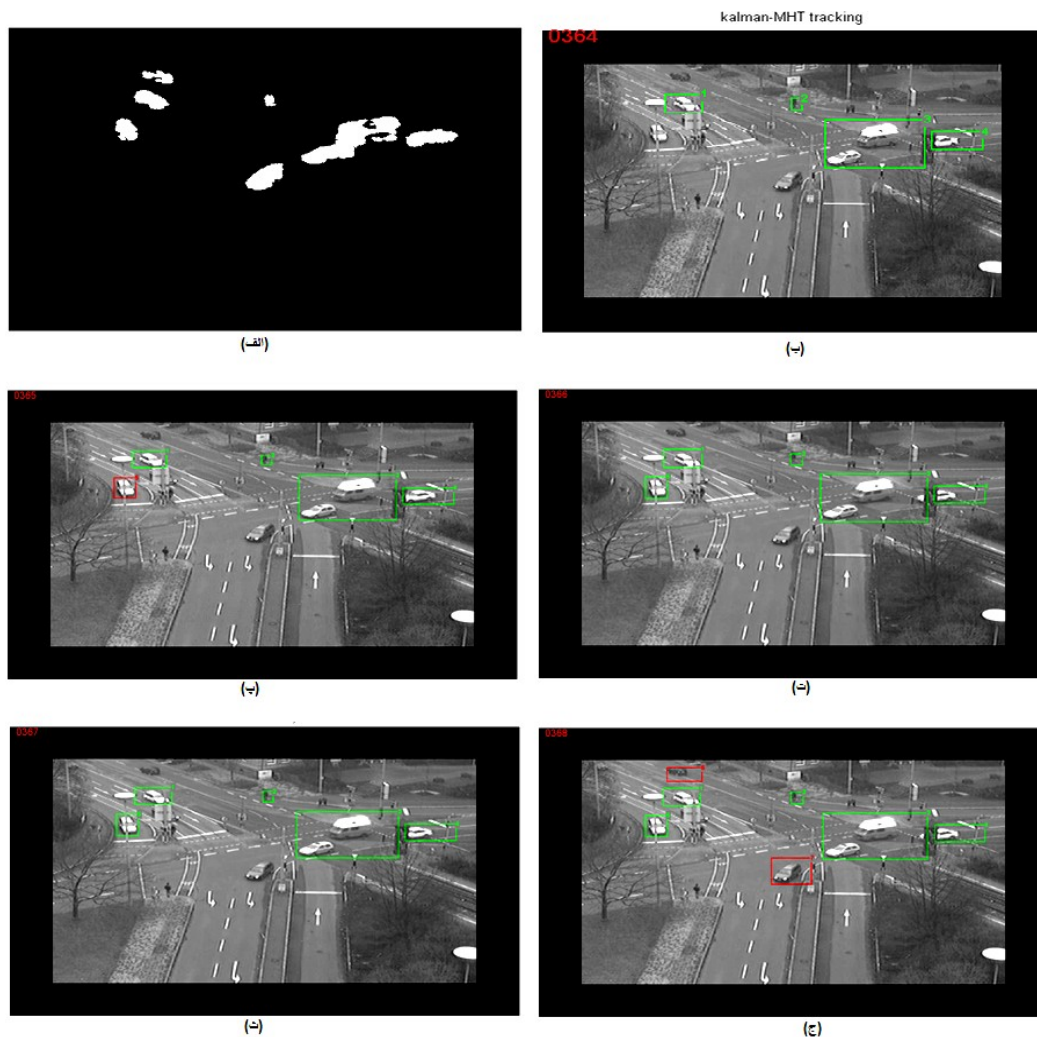
همان‌طور که در تصویر زیر دیده می‌شود، موتور سوار در خروجی سیستم پیشنهادی، حذف شده است.



شکل ۴-۱۸: ردیابی با اعمال سیستم پیشنهادی در دو فریم متوالی

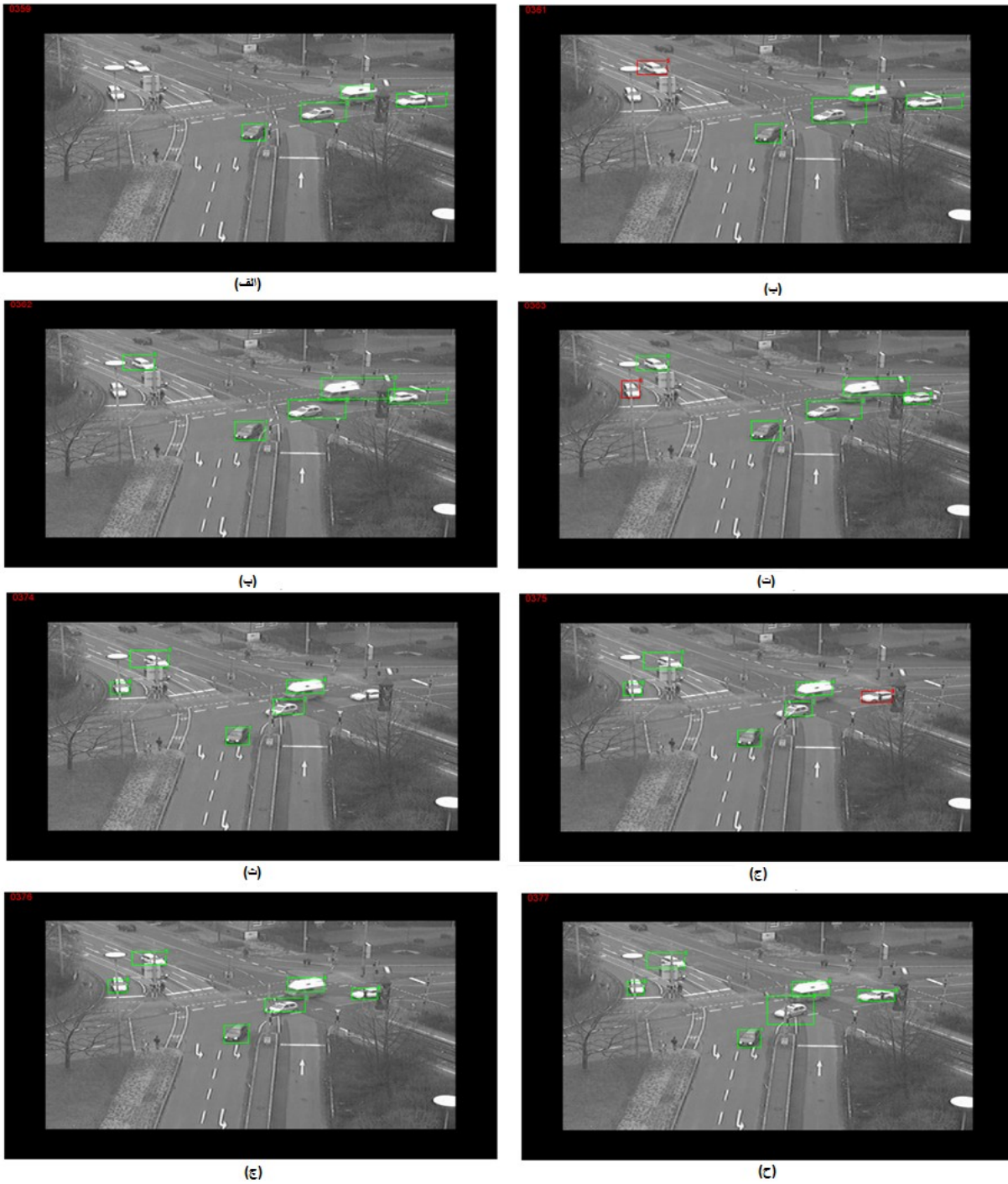
در ادامه حالتی را بررسی می‌کنیم که در آن، علاوه بر غیرخودرو، یکی شدن بلاب‌ها، با استفاده از

سیستم پیشنهادی، رفع می‌گردد.



شکل ۴-۱۹: (الف) خروجی GMM (ب تا ج) ردیابی بدون در نظر گرفتن سیستم با یادگیری و در نتیجه تشخیص و ردیابی ۲ خودرو بعنوان یک خودرو

همان طور که در تصاویر فوق دیده می شود، بر اثر تداخل و یکی شدن بلابها، سیستم ردیابی، دو خودرو را به عنوان یک خودرو در نظر می گیرد. در سیستم پیشنهادی همان طور که در بخش های قبلی توضیح داده شد، این مشکل رفع می گردد. در ادامه خروجی مربوط به سیستم پیشنهادی نشان داده شده است:

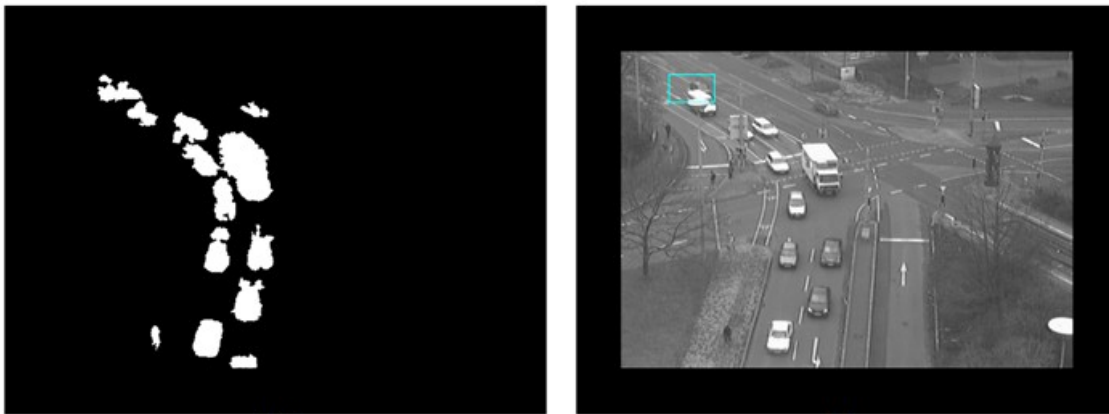


شکل ۴-۲۰: ردیابی با اعمال سیستم پیشنهادی (الف تا ح) نتیجه تشخیص و ردیابی جداگانه خودروها در فریم‌های متوالی

در ادامه به بررسی سیستم تشخیص همپوشانی می‌پردازیم، و در تصاویری که نشان داده خواهد شد، ابتدا سیستم ردیابی بدون در نظر گرفتن الگوریتم پیشنهادی مورد بررسی قرار گرفته، و در ادامه خروجی سیستم تشخیص همپوشانی و نتیجه ردیابی پس از اعمال الگوریتم پیشنهادی نشان داده خواهد شد:



شکل ۴-۲۱: ردیابی با در نظر نگرفتن سیستم پیشنهادی و عدم تشخیص همپوشانی در فریم‌های متوالی



شکل ۴-۲۲: تشخیص همپوشانی توسط سیستم پیشنهادی

خروجی سیستم همپوشانی در شکل ۴-۲۲ نمایش داده شده است، و در نهایت خروجی مربوط به ردیابی همراه با سیستم پیشنهادی در شکل زیر نشان داده می‌شود:



شکل ۴-۲۳: (الف تا پ) ردیابی با اعمال سیستم پیشنهادی و تشخیص همپوشانی

در ادامه خروجی مربوط به تاثیر بخش همپوشانی برای ویدئوهای مختلف، مانند شکل بالا نمایش داده خواهد شد. همانطور که در شکل زیر دیده می‌شود، همپوشانی مربوط به بلاک شماره ۵ تشخیص داده نشده است. در این ویدئو، دو خودرو با هم همپوشانی دارند و خودروی سوم، در فریم‌های مورد نظر حرکت نمی‌کند. در ادامه این همپوشانی توسط واحد مربوطه تشخیص داده می‌شود.



شکل ۴-۲۴: (الف و ب) عدم تشخیص همپوشانی

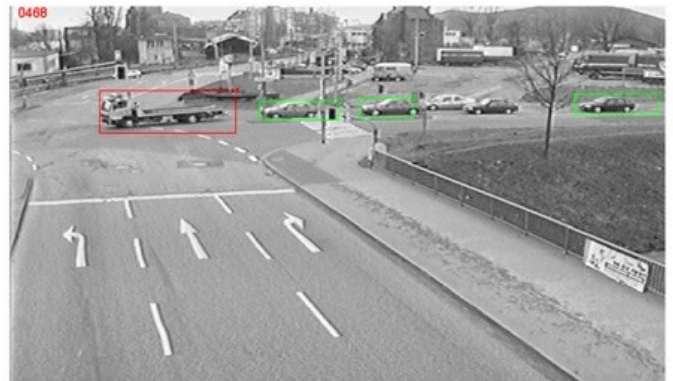


شکل ۴-۲۵: تشخیص همپوشانی

در نهایت ردیابی را برای دو خودرو در قسمت همپوشانی خواهیم داشت:



(الف)



(ب)



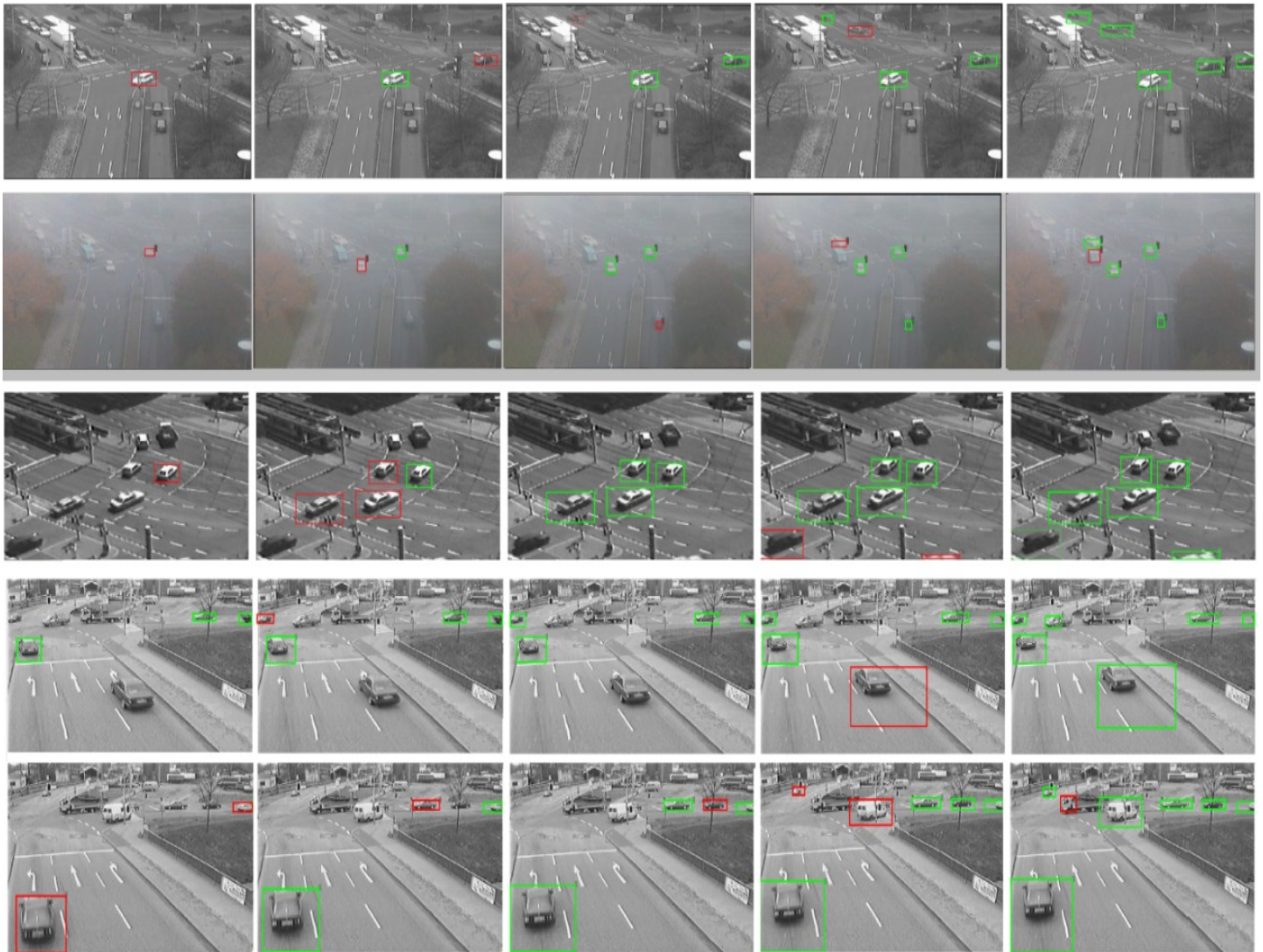
(ج)



(ت)

شکل ۴-۲۶: (الف تا ت) نتیجه ردیابی در چند فریم متوالی با تشخیص همپوشانی

در ادامه نمونه تصاویر مربوط به خروجی نهایی سیستم، برای ۵ ویدئوی مختلف نشان داده خواهد شد.



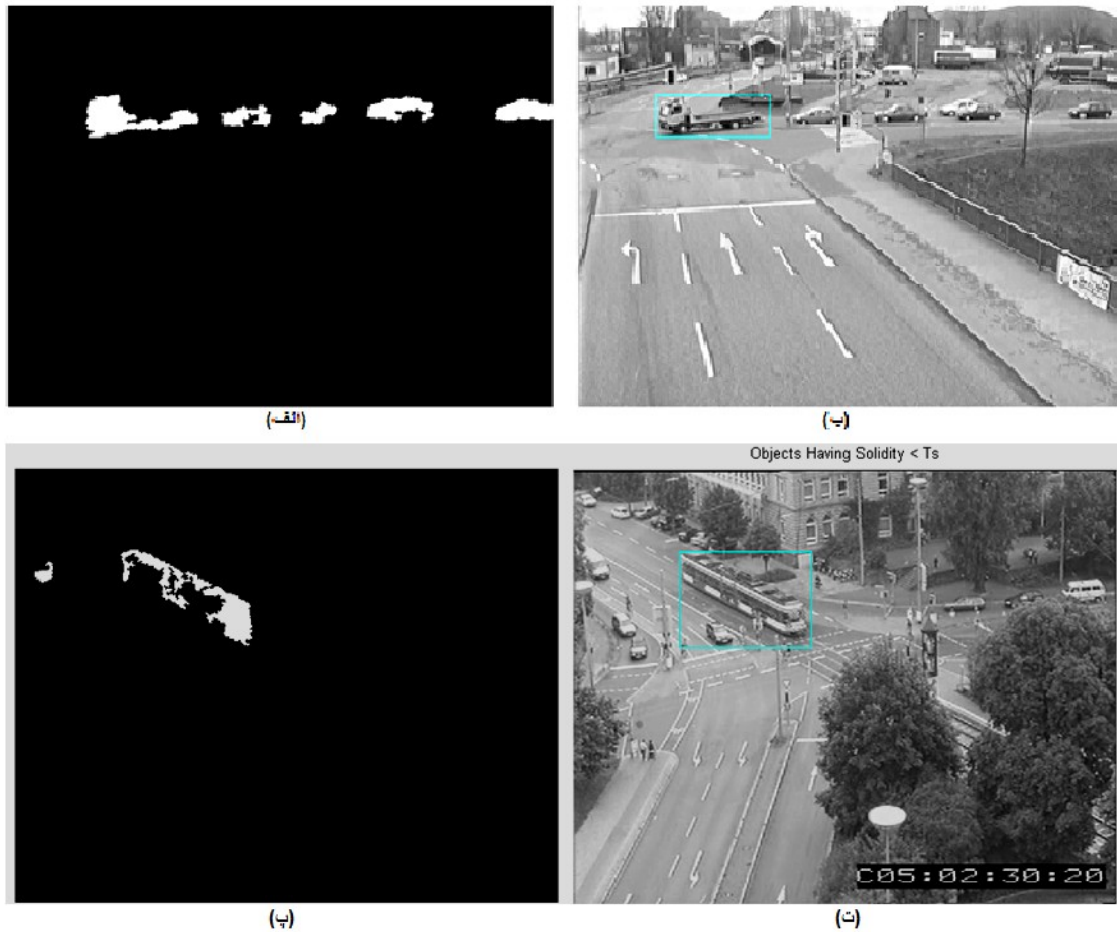
شکل ۴-۲۷: ردیابی نهایی توسط سیستم پیشنهادی در ویدئوهای مختلف

همانگونه که در شکل فوق مشاهده می‌شود الگوریتم پیشنهادی نتایج بسیار مناسبی برای شرایط

محیطی مختلف دارد.

سیستم فوق در برخی از حالات، با مشکلاتی نیز مواجه است که در ادامه مورد بررسی قرار

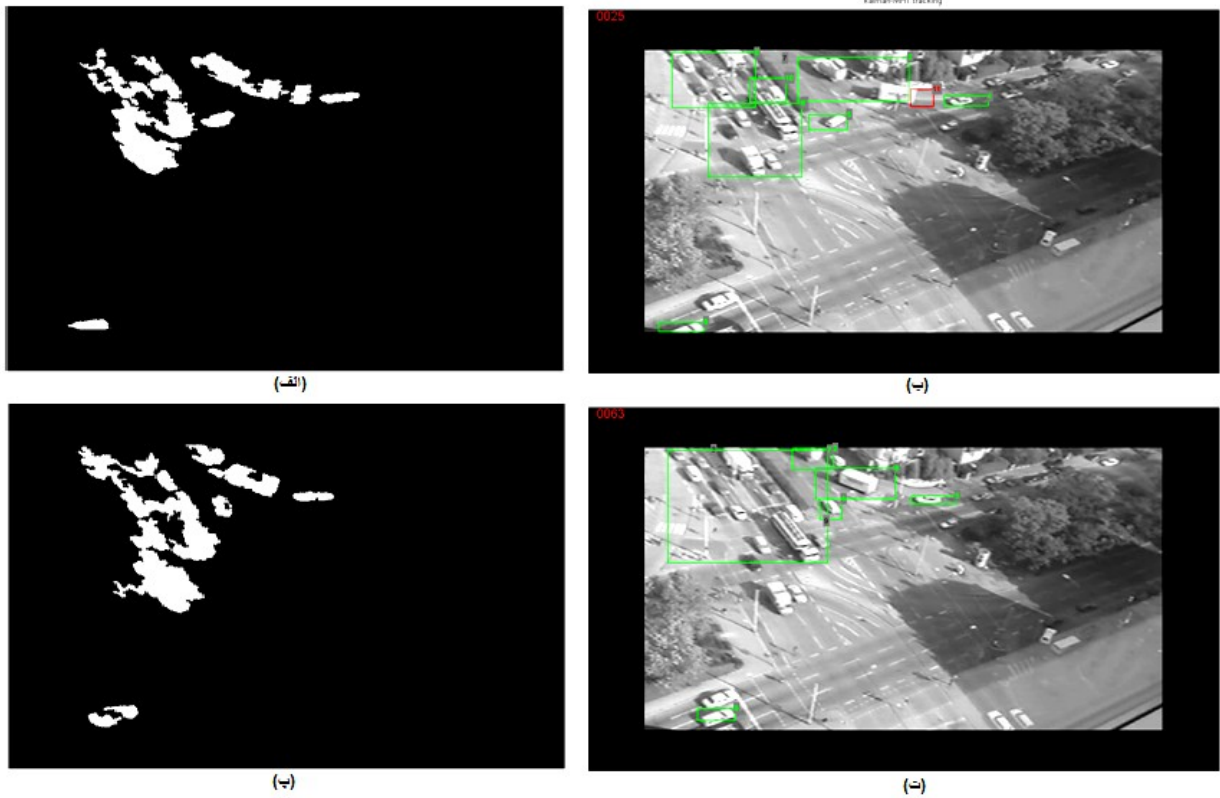
می گیرند.



شکل ۴-۲۸: تشخیص نادرست همپوشانی

در تصاویر فوق، کامیونت و اتوبوس، به دلیل فرم خاص بلاک ، همپوشانی تشخیص داده شده اند، که از نقاط ضعف این سیستم به شمار می رود و نیاز به تقویت واحد تشخیص همپوشانی در چنین شرایطی داریم.

کیفیت بسیار پایین برخی از تصاویر، شرایط نویزی، تعداد بسیار بالای وسایل نقلیه و وجود سایه ها از دیگر عواملی است که سیستم پیشنهادی در مواجهه با آنها دارای نقاط ضعف می باشد.



شکل ۴-۲۹: (الف و پ) خروجی GMM (ب و ت) خطای تشخیص و ردیابی در سیستم پیشنهادی همانطور که در تصویر ذیل مشاهده می‌گردد، پیچیدگی تصویر به حدی بالا می‌باشد، که واحد همپوشانی را با خطاهایی مواجه نموده است:



شکل ۴-۳۰: تشخیص نادرست همپوشانی با وجود اعمال سیستم پیشنهادی

در جدول ۳-۴، نتایج پایانی مربوط به ارزیابی کلی سیستم نشان داده شده است.

جدول ۳-۴: نتیجه نهایی مربوط به ارزیابی کلی سیستم

ویدئو	شمارش و تشخیص صحیح خودرو		شمارش چند خودرو بعنوان یک خودرو		منفی اشتباه (FN)		شمارش بیش از یک بار یک خودرو ^۱		مثبت اشتباه (FP)	
	تعداد	درصد	تعداد	درصد	تعداد	درصد	تعداد	درصد	تعداد	درصد
۱	۳۶۷	%۷۹.۴	۵۴	%۱۱.۷	۴۱	%۸.۹	۲۸	%۶.۸	۱۸	%۴.۳
۲	۲۰۲	%۷۵.۹	۳۸	%۱۴.۳	۲۶	%۹.۸	۱۶	%۶.۹	۱۲	%۵.۲
۳	۴۹	%۷۵.۴	۱۲	%۱۸.۵	۴	%۶.۱	۰	%۰	۴	%۷.۵

این جدول برای ۳ ویدئو و پس از گذشت فریم‌های مختلف در هر یک، تنظیم شده است. نتایج با توجه به حجم ترافیک و تعداد فریم‌های بررسی شده برای هر یک از ویدئوها، نشان داده شده است. یکی از مشکلات تنظیم این جدول، عدم وجود اطلاعات ارزیابی برای هر پایگاه داده بود، بنابراین اطلاعات مورد نظر در هر فریم با توجه به حرکت وسایل نقلیه و تشخیص نهایی در خروجی سیستم ردیابی، تهیه شده‌اند. در ادامه توضیحی بر روند جدول فوق خواهیم داشت:

تشخیص درست خودرو^۲، مربوط به خودروهایی است که به درستی تشخیص داده شده‌اند. شمارش چند خودرو بعنوان یک خودرو^۳، در شرایطی رخ می‌دهد، که چند وسیله نقلیه در خروجی به عنوان یک وسیله نقلیه در نظر گرفته شود و در حقیقت واحد تشخیص همپوشانی وظیفه خود را به درستی انجام نداده باشد.

¹ over segmentation

² True match

³ over group

منفی اشتباه^۱، در حالتی به وجود می آید که، یک وسیله نقلیه در یک فریم شمارش نشود.

مثبت اشتباه^۲، در حالتی ایجاد می شود که یک وسیله نقلیه در خروجی واحد ردیابی به طور اشتباه شمارش شود.

جدول شماره ۳-۴، با توجه به تعاریف فوق، تنظیم شده است. بنابر اطلاعات بدست آمده از این جدول، میانگین دقت تشخیص وسیله نقلیه در این روش برابر با ۷۶/۹ % می باشد.

همان طور که مشاهده شد، سیستم پیشنهادی مشکلات زیادی را در مواجهه با کامیون ها و اتوبوس ها دارد. و این وسایل نقلیه معمولاً بیشتر از یک بار شمارش می شوند چراکه واحد همپوشانی آنها را به عنوان خودروهای شامل همپوشانی تشخیص می دهد. با استفاده از روش های مبتنی بر مدل، می توان تا حدودی بر مشکل فوق غلبه نمود. روش دیگر برای غلبه بر این مشکل، این است که فرض شود که اگر مساحت وسیله شناسایی شده از یک آستانه تجربی بیشتر بود، آن را به عنوان کامیون یا اتوبوس تشخیص دهیم. در کل خطاهای شمارش بیش از یک بار یک خودرو و شمارش چند خودرو بعنوان یک خودرو جدی تر از خطاهای FN,FP می باشند. در تنظیم جدول فوق درصد های مربوط به سه ستون اول، با توجه به مقدار هر بخش تقسیم بر مجموع این سه قسمت، بدست آمده است.

در مورد دو ستون بعدی، مقدار هر بخش با توجه به، اندازه همان بخش، تقسیم بر مجموع آنها با تعداد خودروهایی که بدرستی تشخیص داده شده اند بدست می آید.

¹ false negative

² false positive

۴-۵ - نتیجه‌گیری فصل

عملیات ردیابی در ترافیک سنگین به دلیل انسداد و همپوشانی خودروها و روی هم افتادگی سایه‌های خودروها بر روی یکدیگر، با مشکلات زیادی مواجه است. همچنین پیچیدگی ترافیکی در تقاطع‌ها به دلیل وجود چندین مسیر ورودی و خروجی، زیادتر می‌باشد. وجود چندین جریان ترافیکی و تغییر مسیر حرکت خودروها، وجود عابر پیاده و موتور سیکلت و خودروهای متوقف در تقاطع‌ها، بر پیچیدگی کنترل اتوماتیک ترافیک در تقاطع‌ها و جاده‌های درون شهری، می‌افزاید.

در این پایان‌نامه تلاش شد تا مشکلاتی از قبیل همپوشانی خودروها، تغییرات روشنایی محیط، کیفیت نامناسب تصاویر دریافتی از دوربین‌های ترافیکی، لرزش‌های دوربین، وجود سایه که موجب به‌هم‌پیوستگی خودروها می‌شود، به حداقل رسانیده شوند. سیستم پیشنهادی ما می‌تواند، مشکلاتی از قبیل همپوشانی جزئی و تشخیص خودروها در صحنه‌های پیچیده را برطرف نماید. پیاده‌سازی‌ها با استفاده از پایگاه داده‌های مختلف در شرایط گوناگون جوی از تقاطع‌ها، مورد بررسی قرار گرفته است. خروجی نهایی سیستم فوق عملکرد مناسبی را با میانگین درصد تشخیص ۷۶/۹٪ ارائه داده است.

فصل پنجم

نتیجه‌گیری و پیشنهادات

۵-۱ - نتیجه‌گیری

عملیات ردیابی در ترافیک سنگین به دلیل انسداد و همپوشانی خودروها و روی هم افتادگی سایه‌های خودروها بر روی یکدیگر، با مشکلات زیادی مواجه است. همچنین پیچیدگی ترافیکی در تقاطع‌ها به دلیل وجود چندین مسیر ورودی و خروجی، زیادتر می‌باشد. وجود چندین جریان ترافیکی و تغییر مسیر حرکت خودروها، وجود عابر پیاده و موتور سیکلت و خودروهای متوقف در تقاطع‌ها، بر پیچیدگی کنترل اتوماتیک ترافیک در تقاطع‌ها و جاده‌های درون شهری، می‌افزاید.

در این پایان‌نامه روشی جدید برای به حداقل رساندن خطاهای ناشی از تشخیص و ردیابی خودروها در تقاطع‌ها و صحنه‌های شلوغ، ارائه شده است. در روش پیشنهادی بمنظور تشخیص وسایل نقلیه متحرک از مدل ترکیبی گوسی جهت استخراج اشیاء متحرک موجود در تصویر ورودی استفاده

شد و در ادامه برای حذف عابران پیاده، موتور سیکلت‌ها، خودروهای متوقف در تقاطع‌ها و سایر اشیاء متحرک از خودروهای شناسایی شده و همچنین تقویت خروجی مرحله تشخیص از روش ماشین بردار پشتیبان، که به عنوان قوی‌ترین الگوریتم طبقه‌بندی مطرح می‌باشد استفاده نمودیم. برای ورودی ماشین بردار پشتیبان، به بردار ویژگی بسیار قدرتمندی برای تمایز خودرو از غیرخودرو نیاز داریم که نسبت به جهت و اندازه خودرو مستقل باشد. به همین منظور از توصیف‌گر هیستوگرام گرادیان‌های جهت‌دار استفاده نمودیم.

برای رفع مشکل همپوشانی خودروها، سیستمی جهت تشخیص همپوشانی پیشنهاد شده است که با استفاده از شکل ظاهری و مشخصات مورفولوژیکی خودروها، مکان همپوشانی را تشخیص داده و بمنظور جداسازی خودروهای داخل ناحیه همپوشانی، از یک مدل SVM که با استفاده از تصاویری با اندازه‌های گوناگون آموزش دیده‌است، استفاده می‌نماید. سپس خروجی این مرحله به واحد ردیابی که بر پایه فیلتر کالمن و الگوریتم ردیابی فرضیات چندگانه استوار است وارد شد. علت انتخاب فیلتر کالمن همانگونه که در فصل سوم بطور مفصل شرح داده شد این است که در تخمین و ردیابی اجسامی که حرکتی خطی دارند، بهینه‌ترین پاسخ را می‌دهد. از طرفی فیلتر ذره‌ای با افزایش تعداد وسایل نقلیه متحرک، بسیار کند عمل می‌نماید و از این رو، در یک سیستم بلادرنگ استفاده چندانی ندارد. در نتیجه با توجه به تعداد بالای خودروهای موجود در تقاطع‌ها، و اینکه مسیر حرکتی خودروها تقریباً خطی می‌باشد و پیچیدگی بسیار بالایی در حرکت آنها مشاهده نمی‌شود، از روش فیلتر کالمن به‌مراه تئوری ردیابی فرضیات چندگانه برای ردیابی همزمان تعداد بالایی از وسایل نقلیه متحرک استفاده نمودیم.

سیستم پیشنهادی ما توانست، مشکلاتی از قبیل همپوشانی جزئی و تشخیص خودروها در صحنه‌های پیچیده را برطرف نماید. پیاده‌سازی‌ها با استفاده از پایگاه داده‌های مختلف در شرایط

گوناگون جوی از تقاطع‌ها، مورد بررسی قرار گرفته است. خروجی نهایی سیستم فوق عملکرد مناسبی را با میانگین درصد تشخیص ۷۶/۹٪ ارائه داده است.

۲-۵ - پیشنهادات برای کارهای آینده

استخراج، ثبت و ردیابی خط سیرهای تمامی وسایل نقلیه متحرک در تقاطع‌ها و در نهایت استخراج نقشه‌ی گرمایی بمنظور شناخت مسیرهای پرتردد جهت تقویت سیستم‌های کنترلی در آن مسیرها می‌تواند بعنوان یکی از مهمترین گام‌ها برای کارهای آینده باشد.

استفاده از توصیفگر Hog بصورت ۳ بعدی، با استفاده از روش‌های مبتنی بر مدل کارایی و دقت این سیستم را چندین برابر می‌نماید. بنابراین روش 3DHOG بعنوان گام دیگری برای کارهای آینده پیشنهاد می‌شود.

کارهای موجود در زمینه‌ی ردیابی وسایل نقلیه متحرک، از پایگاه داده‌ی مخصوص به خود استفاده می‌نمایند. دلیل این امر عدم وجود پایگاه داده‌ی استاندارد در این زمینه می‌باشد. به دلیل استفاده از داده‌های مختلف، که قطعاً از نظر شرایط و میزان پیچیدگی متفاوت می‌باشند، مقایسه‌ی کارهای انجام پذیرفته شده را با مشکل مواجه می‌نماید. بنابراین نیاز به یک پایگاه داده‌ی استاندارد به شدت احساس می‌گردد.

واحد همپوشانی در روش پیشنهادی توان تشخیص دقیق همپوشانی‌های جزئی را دارد، اما در مواردی که همپوشانی کلی باشد، سیستم با مشکلاتی مواجه می‌شود. بعنوان مثال سیستم پیشنهادی در مواجهه با کامیون‌ها و اتوبوس‌ها مشکلاتی دارد. و این وسایل نقلیه معمولاً بیشتر از یک بار

شمارش می‌شوند چراکه واحد همپوشانی آن‌ها را به عنوان خودروهای شامل همپوشانی تشخیص می‌دهد. با استفاده از روش‌های مبتنی بر مدل، می‌توان تا حدودی بر مشکل فوق غلبه نمود.



- [1] R. Jain, "Machine Vision", McGraw-Hill, 1995.
- [2] C. Setchell, "Applications of computer vision to Road-traffic Monitoring", PhD thesis, Department of computer science, University of Bristol, September 1997.
- [3] E. Viarani."Extarction of traffic information from image at DEIS"proc of 10th International conference on image Analysis Venice Italy, pp. 1073-1077.Sep 27-29,IEEE Conference Society, 1999
- [4] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," ACM Computing Surveys, vol. 38, no. 4, pp. 1–45, December 2006.
- [5] A. Yilmaz, X. Li and M. Shah, Contour Based Object Tracking with Occlusion Handling in Video Acquired Using Mobile Cameras, IEEE Trans. Patt. Analy. Mach. Intell. 26, 11, pp. 1531-1536.
- [6] S. C. Zhu and A. L. Yuille, "Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation", IEEE Trans. on PAMI vol.18, no.9, pp884-900, 1996
- [7] Paragios,N and Deriche,R.. Geodesic Active Regions for Supervised Texture Segmentation. Int. J.Comput. Vision. pp. 223-247. 2002
- [8] A. Elgammal, R. Duraiswami, D. Harwood and L. S. Davis "Background and Foreground Modeling using Non-parametric Kernel Density Estimation for Visual Surveillance", Proceedings of the IEEE, July 2002.
- [9] D. Comaniciu, V. Ramesh, P. Meer."Kernel-based object tracking."IEEE Trans. Patt.Analy.Mach.Intell.25, 564-575. 2003.

- [10] P. Fieguth, D. Terzopoulos, "Color based tracking of heads and other mobile objects at video frame rates, Proc. Computer Vision and Pattern Recognition Conference (CVPR'97), San Juan, PR, June 1997, 21–27.
- [11] Gareth J. Edwards, Christopher J. Taylor, Timothy F. Cootes. Interpreting Face Images Using Active Appearance Models. In Proceedings of FG'1998. pp.300~305
- [12] MUGHADAM, B. AND PENTLAND, A. Probabilistic visual learning for object representation. IEEE Trans. Patt. Analy.1997. Mach. Intell. 19, 7,pp. 696-710
- [13] M. J. Black and A. D. Jepson, "EigenTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation,"International Journal of Computer Vision, vol. 26, no. 1, pp. 63 – 84, 1998.
- [14] C. Harris and M. Stephens, "A combined corner and edge detector," in Proceedings of The Fourth Alvey Vision Conference, pp. 147-151, 1988.
- [15] David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60, 2. 2004, pp. 91-110.
- [16] H. A. Rowley, Neural Network Based Face Detection, Neural network Based Face Detection. IEEE Trans. Patt. Analy.Mach. Intell. 20,1,23-38.
- [17] P. Viola, Michael J. Jones. D. Snow "Detecting Pedestrians Using Patterns of Motion and Appearance."IEEE International Conference on Computer Vision (ICCV). 734-741.
- [18] Lynne L. Grewe, Avinash C. Kak. Interactive Learning of a Multiple-Attribute Hash Table Classifier for Fast Object Recognition. Computer Vision and Image Understand. 61, 3,387-416.1995.
- [19] S. Avidan. Support Vector Tracking. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2001), Hawaii, December 2001.184-191.
- [20] 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.

- [21] S. Gil, R. Milanese, and T. Pun, “ Feature Selection for Object Tracking in Traffic Scenes.” Intl. Symp. On intelligent Vehicle Highway Systems, society for photo – optical Instrumentation and Engineering, Boston, MA, 1994.
- [22] 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.
- [23] Z. Tang, Z. Miao, and Y. Wan, “Background Subtraction Using Running Gaussian Average and Frame Difference,” Entertainment Computing–ICEC 2007, p. 411–414.
- [24] J. Rymel, J. Renno, D. Greenhill, J. Orwell, and G. a Jones, “Adaptive eigen-backgrounds for object detection,” International Conference on Image Processing. ICIP , 2004, pp. 1847-1850.
- [25] C. Stauffer and W.E.L. Grimson, “Adaptive background mixture models for real-time tracking,” Computer Vision and Pattern Recognition, IEEE Computer Society Conference on. IEEE, 2002, pp. 246-252
- [26] B. Horn, B. Schunk, “Determining optical flow” Artificial intelligent. 17.pp 185-203, 1981.
- [27] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," International Joint Conference on Articial Intelligence, pp. 674-679, 1981.
- [28] M. Black, P. Anandan, “The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields.”, Computer vision image understand. 63, 1, pp 75 – 104. 1996.
- [29] R. Szeliski and J. Coughlan. “Spline-based image registration.” International Journal of Computer Vision, 22(3):199-218, March/April 1997.
- [30] J.L. Barron1, D.J. Fleet and S. Beauchemin, “Performance of Optical Flow Techniques.” IJCV 12:1, pp 43-77, 1994.
- [31] D.koller, K.weber, and H.-H.nage, “Model based object tracking in monocular image sequences of road traffic scenes”.International Journal of Computer Vision 10:3, 1993.

- [32] Manual of “ Open Source Computer Vision Library”, Intel, 2000.
- [33] S. Gil, R. Milanese, and T. Pun, “Feature selection for Object tracking in traffic scenes.” Intl. Symp. On Intelligent Vehicle Highway Systems, society for photo-optical Instrumentation and engineering, vol.2421, pp.173-180, 1994.
- [34] J.M. Roberts, “attentive visual tracking and trajectory estimation for dynamic scene segmenta.” Phd thesis, university of Southampton, 1994.
- [35] Z. Sun, G. Bebis, and R. Miller,” On-Road Vehicle Detection: A Review”. IEEE Transactions on Pattern analysis and machine intelligence, vol. 28, no. 5, MAY 2006.
- [36] Z. Kim and R. Nevatia, “Uncertain Reasoning and Learning for Feature Grouping”, Computer Vision and Image Understanding, vol.76, no.3, pp. 278-288, 1999.
- [37] J. Malik and S. Russell, “Traffic Surveillance and Detection Technology Development: New Traffic Sensor Technology”, PATH research report UCB-ITS-PRR-97-6, Institute of transportation studies, University of California, Berkeley, 1997.
- [38] J. Badenas, M. Sanchiz, Filiberto Pla, “Using Temporal Integration for Tracking Regions in Traffic Monitoring Sequences.” ICPR: 7137-7140. 2000.
- [39] I. Pavlidis, V. Morellas, P. Tsiamyrtzis, and S. Harp, "Urban Surveillance System: from the Laboratory to the Commercial World," Proc. IEEE, vol. 89, pp.1478-1497, Oct, 2001.
- [40] T. Ellis, "Multi Camera Video Surveillance," Proc. 36th Annual 2002 Int. Carnahan Conf. on Security Technology, Atlantic City, New Jersey, pp. 228-233,2002.
- [41] Ho C.Y., Lee R.C.K., "A Bayesian approach to problems in stochastic estimation and control " IEEE Transactions on Auto.Control, 1964, Vol. AC (9), pp.333-339.
- [42] G. Welch and G. Bishop, “An introduction to the Kalman filter,” University of North Carolina at Chapel Hill, Chapel Hill, NC, 1995.

- [43] M.I. Ribeiro, “Kalman and extended kalman filters: Concept, derivation and properties,” Institute for Systems and Robotics, Lisboa Portugal, 2004.
- [44] Blackman, S.S. Multiple hypothesis tracking for multiple target tracking aerospace and Electronic Systems Magazine, IEEE Jan. 2004 Volume: 19 Issue: 1 pp: 5 – 18.
- [45] T. E. Fortman, Y. Bar-Shalom, M. Scheffe, “Sonar Tracking of Multiple Targets Using Joint Probabilistic Data Association,” IEEE Journal of Oceanic Engineering, Vol. OE-8, pp. 173-184, 1983.
- [46] 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.
- [47] Isard M., Blake A., "ICondensation: Unifying low-level and high-level tracking in a stochastic framework", In Proceeding of European Conference on Computer Vision, volume 1, pp. 767-781, Freiburg, Germany, June1998.
- [48] A. Ticiano, C. Bragatto., “A new approach to multiple vehicle tracking in intersections using harris corners and adaptive background subtraction” IEEE Intelligent Vehicles Symposium, Eindhoven University of Technology Eindhoven, The Netherlands, June 4-6, 2008.
- [49] A. Cavallaro, E. Salvador, T. Ebrahimi, and Q. Mary, “DETECTING SHADOWS IN IMAGE SEQUENCES,” 2004, pp. 165-174.
- [50] C. Wang and W. Zhang, “A Robust Algorithm for Shadow Removal of Foreground Detection in Video Surveillance,” Asia-Pacific Conference on Information Processing, Jul. 2009, pp. 422-425.
- [51] Dalal, N. Triggs, B.: Histograms of Oriented Gradients for Human Detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.
- [52] T. Tisdall and Yamada, “Review and MATLAB Implementation of “Histogram of Oriented Gradients for Human Detection” by Dalal and Triggs”.

-
- [53] B. Tamersoy and J.K. Aggarwal, "Robust Vehicle Detection for Tracking in Highway Surveillance Videos using Unsupervised Learning" Advanced Video and Signal Based Surveillance, Computer and Vision Research Center The University of Texas at Austin, 2009.
- [54] G. Jun, J. K. Aggarwal, M. Gökmen, "Tracking and Segmentation of Highway Vehicles in Cluttered and Crowded Scenes", WACV '08: IEEE 2008 Workshops on Applications of Computer Vision Copper, Colorado, Jan. 2008 (WACV '08)
- [55] <http://i21www.ira.uka.de/>
- [56] i-LIDS dataset for AVSS 2007.
- [57] Z. Zivkovic."Improved adaptive Gaussian mixture model for background subtraction", International Conference Pattern Recognition, Vol.2, 2004 pp:28-31.
- [58] <http://svmlight.joachims.org/>

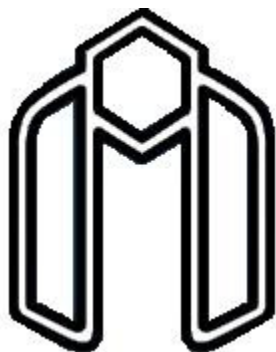
Abstract:

One of the most important research efforts in Intelligent Transportation Systems (ITS) is the development of systems that automatically monitor the flow of traffic at intersections. Rather than being based on aggregated flow analysis, these systems should provide detailed data about each vehicle, such as its position and speed in time. These systems would be useful in reducing the workload of human operators, in improving our understanding of traffic and in alleviating such dire problems as congestion and collisions that plague the road networks.

In this research, a novel approach to vehicle detection and tracking in intersections and urban roads where it is difficult to track individual vehicles in heavy traffic because vehicles either occluded each other or are connected together by shadows. Scenes from traffic monitoring are usually noisy due to weather condition. We use background subtraction based on Gaussian mixture model to extract the foreground. For accurate object detection, an efficient scheme is used to remove shadows and noisy speckles. Because of some difficulties mentioned above, we used a set of training image which have captured automatically, to train a classifier. Decision boundary is computed using Support Vector machines (SVMs). A Histogram of Oriented Gradient (HOG) as a feature descriptor is used to train an SVM classifier. An occlusion system detects blobs that are suspected of having more than one vehicle in them by morphological analysis. Finally, for tracking the vehicles we used a Kalman filter with a multiple hypothesis tracking algorithm. The proposed system has demonstrated a good performance for intersections and urban roads video sequence with an average percentage of correctly detected vehicles of 76.9 %.

Keywords:

Vehicle detection and tracking; background subtraction; Gaussian Mixture Model; Histogram of Oriented Gradients; Support Vector Machines; kalman filter, Multiple Hypothesis Tracking.



Shahrood University of Technology

Faculty of Electrical and Robotics Engineering.

Vehicle Detection and Tracking

Using prediction methods

Salehe Zohroudi

Supervisor(s):

Dr. Omid Reza Marouzi

Dr. Alireza Ahmadifard

Feburary 2012

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.