

## جبـانـ حـرـكـتـ دـورـبـينـهـایـ كـنـترـلـ نـظـارـتـیـ باـ استـفـادـهـ اـزـ ثـبـتـ تصـاوـيرـ متـواـليـ درـ شـنـاسـايـ خـودـروـهاـ

علـیـ کـرمـیـ<sup>۱</sup>، مـحـسـنـ سـرـیـانـیـ<sup>۲</sup>، مـسـعـودـ وـرـشـوـسـازـ<sup>۳</sup>

۱- دـانـشـجوـیـ کـارـشـناسـیـ اـرـشـدـ فـتوـگـرامـتـیـ، دـانـشـکـدـهـ نقـشـهـ بـرـدـارـیـ، دـانـشـگـاهـ صـنـعـتـیـ خـواـجـهـ نـصـیرـالـدـینـ طـوـسـیـ

۲- دـانـشـیـارـ گـروـهـ هـوشـ مـصـنـعـیـ، دـانـشـکـدـهـ مـهـنـدـسـیـ کـامـپـیـوتـرـ، دـانـشـگـاهـ عـلـمـ وـ صـنـعـتـ اـیـرانـ

۳- دـانـشـیـارـ گـروـهـ فـتوـگـرامـتـیـ وـ سـنجـشـ اـزـ دـورـ، دـانـشـکـدـهـ نقـشـهـ بـرـدـارـیـ، دـانـشـگـاهـ صـنـعـتـیـ خـواـجـهـ نـصـیرـالـدـینـ طـوـسـیـ

تـارـیـخـ دـرـیـافتـ مـقـالـهـ: ۱۳۹۵/۱۰/۲۸ تـارـیـخـ پـذـیرـشـ مـقـالـهـ: ۱۳۹۵/۱۰/۲۸

### چـکـيـدهـ

تشـخـيـصـ خـودـروـهـاـ درـ تصـاوـيرـ وـيـدـئـوـيـ بـهـعـنـوانـ يـكـ مـوـضـوعـ كـلـيـدـيـ درـ مـبـاحـثـ بـيـنـايـيـ کـامـپـيـوتـرـ مـحـسـوبـ مـيـگـرـددـ. فـرـضـ اـصـلـيـ اـكـثـرـ دـورـبـينـهـایـ نـظـارـتـيـ، ثـبـتـ بـودـنـ دـورـبـينـ درـ حـيـنـ فـيلـمـ بـرـدـارـيـ مـيـيـاشـدـ وـلـيـ اـنـگـرـهـ نـظـارـتـيـ بـهـ هـرـ دـلـيلـيـ مـثـلاـ بـرـاـزـ بـادـ يـاـ فـشـارـهـاـيـ بـيـرونـيـ مـحـيـطـ اـطـرافـ دـچـارـ لـرـزـشـهـاـيـ مـكـرـرـ شـوـدـ دـيـگـرـ رـوـشـهـاـيـ مـوـجـودـ بـرـايـ شـنـاسـايـ خـودـروـهـاـ بـهـ دـرـستـيـ عـلـمـ نـمـيـكـنـدـ. اـيـنـ مـوـضـوعـ باـعـثـ جـابـجاـ شـدـنـ پـيـكـسـلـهـاـيـ تـصـوـيرـ وـيـجادـ نـويـزـ مـيـشـوـدـ كـهـ دـرـ اـكـثـرـ مـوـاقـعـ باـعـثـ شـنـاسـايـ عـوـارـضـ ثـابـتـ بـهـعـنـوانـ عـوـارـضـ مـتـحـركـ مـيـگـرـددـ. هـدـفـ اـصـلـيـ درـ اـيـنـ تـحـقـيقـ اـسـتـفـادـهـ اـزـ رـوـشـ ثـبـتـ (ـهـمـمـرـجـعـ سـازـيـ)ـ فـريـمـهـاـيـ مـتـوـالـيـ بـاـ تصـوـيرـ پـسـ زـمـينـهـ، بـرـايـ حـذـفـ نـويـزـ وـ مـشـكـلـ يـادـشـدهـ مـيـيـاشـدـ. روـنـدـ كـارـ بـهـ اـيـنـ صـورـتـ مـيـيـاشـدـ كـهـ بـعـدـ اـزـ اـسـتـخـرـاجـ فـريـمـهـاـيـ مـتـوـالـيـ وـ تصـوـيرـ پـسـ زـمـينـهـ اـزـ طـرـيقـ رـوـشـ مـيـانـهـگـيرـيـ، تـامـ فـريـمـهـاـيـ مـوـجـودـ بـاـ تصـوـيرـ پـسـ زـمـينـهـ ثـبـتـ شـدـهـ وـ سـپـسـ بـاـ اـسـتـفـادـهـ اـزـ تـغـرـيقـ پـسـ زـمـينـهـ بـهـ شـنـاسـايـ خـودـروـهـاـ پـرـداـختـهـ مـيـشـوـدـ. سـهـ نوعـ دـادـهـ بـرـايـ پـيـادـهـسـازـيـ وـ اـرـزـيـابـيـ توـانـايـيـ رـوـشـ پـيـشـنـهـادـيـ مـورـدـ اـسـتـفـادـهـ قـرـارـ گـرفـتـهـ اـسـتـ. وـيـشـگـيـ اـصـلـيـ اـيـنـ مـجـمـوعـهـ دـادـهـاـ لـرـزـشـهـاـيـ مـكـرـرـ دـورـبـينـ درـ حـيـنـ فـيلـمـ بـرـدـارـيـ وـ مـتـفـاقـوـتـ بـودـنـ تـراـكـمـ خـودـروـهـاـيـ مـوـجـودـ درـ هـرـ مـجـمـوعـهـ مـيـيـاشـدـ. اـزـ مـعـيارـهـاـيـ HRـ (ـنـرـخـ تـشـخـيـصـ درـسـتـ)ـ MOPDـ (ـدـقـتـ كـشـفـ چـنـدـ شـيـئـ)ـ، FARـ (ـنـرـخـ تـشـخـيـصـ اـشـتـبـاهـ)ـ وـ OAـ (ـشـاخـصـ دـقـتـ كـلـيـ)ـ بـرـايـ اـرـزـيـابـيـ رـوـشـ پـيـشـنـهـادـيـ درـ دـوـ حـالـتـ بـاـ ثـبـتـ تصـوـيرـ(ـرـوـشـ پـيـشـنـهـادـيـ)ـ وـ بـدـونـ ثـبـتـ تصـوـيرـ استـفـادـهـ شـدـهـ اـسـتـ. بـهـ طـورـ كـلـيـ دـقـتـ بـهـ دـسـتـآـمـدـهـ بـرـايـ شـنـاسـايـ خـودـروـهـاـ بـاـ ثـبـتـ تصـاوـيرـ وـ بـدـونـ ثـبـتـ تصـاوـيرـ بـهـ طـورـ مـتوـسـطـ بـهـ تـرـتـيـبـ ۷۶ـ درـصـدـ وـ ۷۹ـ درـصـدـ مـيـيـاشـدـ. بـنـاـرـايـنـ مـيـيـاشـدـ. اـزـ رـوـشـ ثـبـتـ تصـاوـيرـ حدـودـ ۱۳ـ درـصـدـ دـقـتـ شـنـاسـايـ خـودـروـهـاـ رـاـ اـفـزـاـيشـ دـادـ.

كـلـيـدـ وـاـزـهـهـاـ: ثـبـتـ تصـاوـيرـ، شـنـاسـايـ عـوـارـضـ، تـنـاظـرـيـابـيـ، عـوـارـضـ مـتـحـركـ.

\* نـوـيـسـنـدـهـ مـكـاتـبـهـ كـنـنـدـهـ: تـهـرـانـ - دـانـشـكـدـهـ نقـشـهـ بـرـدـارـيـ - دـانـشـگـاهـ صـنـعـتـیـ خـواـجـهـ نـصـیرـالـدـینـ طـوـسـیـ

تـلـفـنـ: ۹۱۸۷۸۳۹۰۸۶

**۱- مقدمه**

اضافی دوربین قابل تشخیص و حذف می‌باشد و همچنین خطای دیلای در شناسایی خودروها اثر گذار نخواهد بود. بنابراین در این مقاله از پس‌زمینه به عنوان سیستم مختصات مرجع استفاده شده است. در مرحله اول تمام فریم‌های جاری را با پس‌زمینه هم مرجع کرده و بعد از طریق تفاضل به شناسایی خودروها پرداخته می‌شود. این امر باعث می‌شود که مسئله تاخیر احتمالی تصویر برداری به طور کامل رفع و هیچ وابستگی به آن نداشته باشد.

ساختار کلی مقاله در ادامه به این صورت است که در بخش دوم مروری بر تحقیقات پیشین خواهیم داشت. در بخش سوم متدولوزی نحوه جبران حرکت تشریج می‌گردد. در بخش چهارم چگونگی تشخیص خودروها بررسی می‌شود. در بخش پنجم به نحوه پیاده‌سازی اشاره شده است. در بخش ششم ارزیابی روش پیشنهادی صورت گرفته است. جمع‌بندی کلی در بخش هفتم صورت گرفته است و نهایتاً در بخش هشتم پیشنهاداتی برای تحقیقات آینده خواهیم داشت.

**۲- پیشینه تحقیق**

تاکنون تحقیقات زیادی به منظور جبران حرکت دوربین صورت گرفته است. از این رو در ادامه تعدادی از روش‌های موجود را معرفی خواهیم کرد. در [۶] برای تشخیص و ردیابی عابرین پیاده با استفاده از یک دوربین متحرک از جبران حرکت برای ثبت فریم‌های متوالی استفاده کردند. برای تشخیص حرکت دوربین از روش تطبیق بلوک<sup>۴</sup> استفاده شد. در این روش یک بلوک را در فریم لحظه  $t$  در نظر گرفته و بهترین تناظر آن را در فریم  $t+1$  بدست آوردن و بردار جابجایی آن بلوک را به عنوان بردار حرکت دوربین درنظر گرفتند. در [۹] ردیابی اشیاء با دوربین تعییه شده روی یک ربات متحرک در محیط‌های بیرونی بررسی شده است. در این کار برای جبران

تشخیص خودروها در تصاویر ویدئویی به عنوان یک موضوع کلیدی در مباحث بینایی کامپیوتر محسوب می‌گردد [۱]. اولین مرحله در سیستم‌های ردیابی، آشکارسازی عوارض متحرک در تصاویر می‌باشد [۲ و ۳]. در سال‌های اخیر، روش‌های گوناگونی به منظور استخراج و تشخیص عوارض متحرک با استفاده از دوربین‌های نظارتی پیشنهاد شده است. فرض اصلی اکثر دوربین‌های نظارتی، ثابت بودن دوربین در حین فیلمبرداری می‌باشد [۳ و ۴]. در چنین حالاتی، اگر دوربین نظارتی به هر دلیلی همانند وزش باد یا فشارهای بیرونی محیط اطراف دچار لرزش‌های مکرر شود دیگر روش‌های موجود برای شناسایی خودروها به درستی عمل نمی‌کنند. دلیل این موضوع آنست که پیکسل‌های تصویر جایه‌جا می‌شوند که این خود در اکثر مواقع باعث شناسایی عوارض ثابت به عنوان عوارض متحرک می‌شود [۱ و ۵].

تاکنون روش‌های مختلفی برای جبران حرکت دوربین نظارتی انجام شده است. به عنوان مثال Lim<sup>۱</sup> (سال‌های ۲۰۰۵ و ۲۰۱۳ در [۶ و ۷] و همچنین Chen<sup>۲</sup> سال ۲۰۱۲ در [۸] روش‌هایی ارائه کرده بودند که در آن با استفاده از تناظریابی و تفاضل بین دو فریم متوالی به مسئله جبران حرکت دوربین پرداخته بود. به طور کلی مشکل اصلی این روش‌ها این است که دقت روش به ساختار داخلی دوربین مانند دیلای (Delay) وابسته می‌باشد. هدف اصلی در این تحقیق استفاده از روشی نوین جهت برطرف کردن مسئله فوق است. برای این منظور اگر بتوان هر دو فریم پس‌زمینه و فریم جاری را به یک سیستم مختصات مشخص انتقال داد (هم‌مرجع کرد) در آن صورت اختلافات ناشی از حرکات

<sup>1</sup> Lim<sup>2</sup> Chen<sup>3</sup>. تاخیر احتمالی در تصویر برداری<sup>4</sup> Block Matching Algorithm

استفاده از یک الگوریتم کشف حرکات پرست به دست آمده و موقعیت جدید ویژگی‌ها با فیلتر کالمن محاسبه می‌شود. در [۱۳ و ۱۴] نیز بر اساس ویژگی گوشه به جبران حرکت دوربین و ردیابی پرداخته‌اند. در [۱۵] روشی برای تخمین حرکت دوربین و تشخیص اشیاء متحرک ارائه شده است. در این روش ابتدا تصویر اولیه به صورت یک شبکه از مثلث‌ها ساختاربندی شده و از رئوس این مثلث‌ها به عنوان ویژگی بین فریم‌های متوالی استفاده می‌شود. سپس در فریم بعدی ویژگی‌های متناظر را پیدا کرده و بر اساس همسایگی این نقاط متناظر شبکه را به روز رسانی می‌کنند. سرانجام با استفاده از این نقاط متناظر پارامترهای حرکت دوربین را به دست آورده و بردارهایی که به پس زمینه تعلق ندارند را به عنوان اشیاء متحرک موجود در صحنه برچسب گذاری خواهند کرد. در [۱۶] برای دو فریم متوالی نقاط ویژگی متناظر را پیدا کرده و با حذف نقاط ویژگی پرست (نقاط ویژگی مرتبط با اشیاء متحرک) حرکت پس زمینه را تشخیص می‌دهند و با اعمال این حرکت به فریم قبلی و در ادامه اجرای الگوریتم تفاضل فریم ناحیه‌های مربوط به اشیاء متحرک را کشف می‌کنند.

همان گونه که در اکثر روش‌های ذکر شده مشاهده می‌شود، برای حل مسئله جبران حرکت دوربین با استفاده از استخراج ویژگی، تناظریابی و تفاضل بین دو فریم متوالی <sup>۱</sup> و <sup>۲</sup>+۱ عمل شده است. اگر دوربین مورد استفاده خطای تاخیر احتمالی داشته باشد، دقت روش‌های یاد شده برای جبران حرکت تحت تاثیر این خطای قرار می‌گیرد. بنابراین هدف اصلی در این تحقیق ارائه روشی جهت این مسئله می‌باشد. برای این کار اگر یک سیستم مختصات مرجع را ثابت در نظر بگیریم و تمام فریم‌های متوالی را با این سیستم ثبت کنیم، هرچقدر که خطای دیلای داشته باشیم هیچ تاثیری بر روی دقت جبران حرکت دوربین ندارد و کاملاً مستقل از این خطای عمل می‌کند. تنها خطایی که باقی می‌ماند تغییر در موقعیت مکانی

حرکت دوربین از تناظر بین مجموعه‌ای از ویژگی‌ها در دو فریم متوالی استفاده شده است.

در [۸ و ۱۰] برای ردیابی اشیاء توسط یک دوربین متحرک از روش جریان نوری<sup>۳</sup> استفاده شده است. نویسنده‌گان در این مقاله از جریان نوری برای کشف حرکت پیکسل‌ها در بین فریم‌های متوالی و به دست آوردن بردار حرکت دوربین از طریق بردار جاگایی با بیشترین تکرار در بین تمام پیکسل‌ها استفاده کرده‌اند. در [۷] کشف و ردیابی چندین عابر پیاده با استفاده از اطلاعات رنگ و الگوریتم ادابوست<sup>۴</sup> ارائه شده است. در این کار نویسنده‌گان با استفاده از تناظر ویژگی‌های گوشه به دست آمده با الگوریتم هریس<sup>۵</sup> بین دو فریم متوالی حرکت دوربین را به دست می‌آورند سپس با اعمال حرکت دوربین به فریم ۱-۱ و تفاضل این فریم از فریم <sup>۶</sup>، ناحیه مربوط به عابرین پیاده مشخص می‌گردد. در [۱۱] ردیابی اشیاء متحرک بر اساس ویژگی با استفاده از یک دوربین دارای حرکات انتقالی<sup>۷</sup>، مورد بررسی قرار می‌گیرد. در این روش نویسنده‌گان از نقاط گوشه که نسبت به تغییرات روشنایی حساسیت کمتری دارند برای ردیابی استفاده کرده‌اند. پس از آنکه حرکت دوربین با استفاده از یک الگوریتم تطبیق الگو برای یافتن بیشترین شباهت بین دو قالب در دو فریم متوالی تخمین زده شد، توانستند از حرکت‌های غیر واقعی حاصل از حرکت دوربین صرفظیر کنند. ایراد اصلی این روش محدود کردن حرکت دوربین به حرکت انتقالی می‌باشد. در [۱۲] نیز از ویژگی‌های گوشه به دست آمده با الگوریتم فست<sup>۸</sup> برای جبران حرکت دوربین<sup>۹</sup> استفاده شد. در این روش موقعیت اشیاء متحرک با

<sup>۱</sup> Optical Flow

<sup>۲</sup> AdaBoost

<sup>۳</sup> Hariss

<sup>۴</sup> Pan Tilt Zoom

<sup>۵</sup> FAST

<sup>۶</sup> Compensating Camera Movements

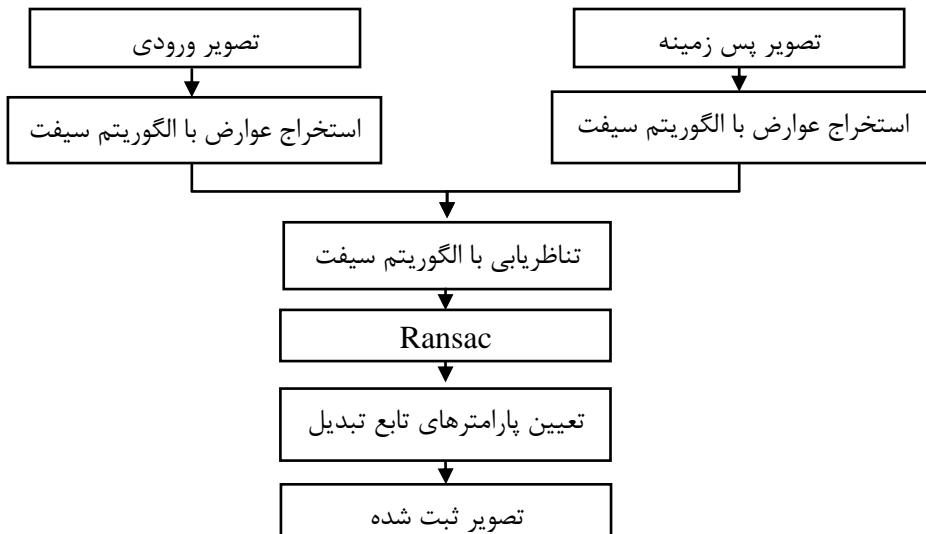
محاسبات داده‌های تصویری از مشکلات اصلی در این زمینه هستند. لذا احتیاج به روش‌هایی که بتوانند راه حل سریع‌تر و ساده‌تری را در این زمینه ارائه کنند ضروری می‌باشد. با توجه به طیف وسیع دوربین‌های موجود، تنوع اعوجاجات و خطاهای موجود در تصاویر، طراحی یک روش عمومی کاربردی در تمامی مسائل ثبت تصاویر غیرممکن می‌باشد. درنتیجه در موارد مختلف با توجه به نوع تصاویر و کاربردهای مورد نظر، باید روشی را طراحی نمود تا بتواند نه تنها خطاهای اعوجاجات هندسی را تصحیح نماید، بلکه نویز تصاویر که بر روی دقت ثبت مؤثر است را نیز در نظر بگیرد. یک روش مناسب می‌تواند با ترکیب انتخاب‌های موجود برای مراحل مختلف فرآیند ثبت که در شکل (۱) توضیح داده شده‌اند، طراحی شود. در ادامه به تشریح مراحل الگوریتم پیشنهادی پرداخته می‌شود.

پیکسل‌ها می‌باشد که با استفاده از ثبت تصاویر ورودی به تصویر پس زمینه (سیستم مختصات مرجع) این خطا هم جبران می‌شود و بهتر به شناسایی خودروها می‌توان پرداخت. در ادامه روش پیشنهادی برای جبران حرکت دوربین تشریح و مورد ارزیابی قرار می‌گیرد.

### ۳- جبران حرکت از طریق ثبت تصاویر به پس زمینه

ثبت تصویر روند تعیین بهترین تطبیق مکانی بین دو یا چند تصویر به دست آمده در زمان‌های مختلف، از منظرهای متفاوت و یا با سنسورهای مختلف می‌باشد که دو تصویر مبنا و ورودی را به صورت هندسی بر هم منطبق می‌نماید [۱۷]. در سال‌های اخیر تحقیقات گسترده‌ای در زمینه تعیین اتوماتیک تابع تبدیل ریاضی برای ثبت تصاویر انجام پذیرفته است [۲۲، ۲۱، ۲۰، ۱۸].

پیچیدگی جستجوی تشابهات، زمان و حجم بالای



شکل ۱: فلوچارت پیشنهادی برای ثبت تصاویر

دو دسته اصلی تقسیم می‌شوند که یا عارضه مبنا و یا ناحیه مبنا هستند [۲۲]. الگوریتم سیفت یکی از روش‌های موجود برای استخراج عوارض متناظر بوده و غالباً برای کاربردهایی نظیر تشخیص شی، تناظریابی تصاویر، ردیابی، بازیابی شی در بانک‌های اطلاعاتی

### ۳-۱- استخراج عوارض

اولین مرحله برای ثبت دو تصویر، استخراج عوارض (ویژگی‌ها) می‌باشد. عارضی که تاکنون برای برقراری ارتباط دو تصویر استفاده گردیده‌اند بسیار متنوع می‌باشند و روش‌های تناظریابی بر این اساس به

الگوریتم رنسک<sup>۱</sup> و در قالب مدل هموگرافی، صحت انطباق جفت تناظرهای اولیه بررسی شده و انطباق‌های غیر صحیح حذف می‌شوند [۲۷]. این الگوریتم توسط فیشلر و بولس<sup>۲</sup> در سال ۱۹۸۱ معرفی شد [۲۷]. الگوریتم رنسک روشی نیرومند به منظور برآشش یک مدل هندسی به مجموعه‌ای از موقعیت‌های متناظر است که تعدادی جفت‌های متناظر اشتباہ نیز در میان آن‌ها وجود دارد [۲۷].

رنسک فرض می‌کند که با مجموعه‌ای از داده‌های صحیح<sup>۳</sup> می‌توان پارامترهای مدل را به گونه‌ای تخمین زد که به طور بهینه به این داده‌ها برآشش شود. داده‌های پرت<sup>۴</sup> از مقادیر زیاد نویز یا از اندازه‌گیری‌های غلط یا فرض‌های نادرست در تفسیر داده رخ می‌دهند و نمی‌توانند به مدل، برآشش شوند. این تخمین با احتمال مشخصی صحیح است، زیرا رنسک یک تخمین زننده اتفاقی می‌باشد. این الگوریتم به طور وسیع در مسائل تخمین پارامترهای مدل در بینایی کامپیوترا از قبیل تناظریابی عوارض، و آشکارسازی هندسی کاربرد دارد.

#### ۳-۴-۳- تعیین پارامترهای تبدیل

با توجه به موقعیت و شرایط فیلم‌برداری که داریم امکان وجود تبدیل‌هایی چون تبدیل افاین و پروجکتیو و ... در بین دو تصویر مبنا و ورودی بالاست. بنابراین در تناظریابی این‌گونه از تصاویر، استفاده از معیار شباهتی که بر پایه‌ی این تبدیل‌ها باشد، دارای توجیه می‌باشد [۲۸].

رابطه (۱) تبدیل افاین مربوط به سطح،  $T: R^2 \rightarrow R^2$  می‌تواند به صورت ماتریس  $2 \times 2$  غیرسینگولار A، و بردار انتقال  $t \in R^2$ ، بیان گردد.

$$\text{رابطه (۱)} \quad p' = A * P + t$$

که  $p \in R^2$  می‌باشد. این تبدیل ترکیبی از انتقال و

چندرسانه‌ای، ربات‌های خودکار محلی و غیره مورد استفاده قرار می‌گیرد. الگوریتم سیفت کارایی بالایی برای تصاویر با تغییر شکل پیچیده و تغییرات روشنایی داشته و نسبت به سایر الگوریتم‌ها سرعت محاسبه و دقت تعیین موقعیت بالاتری دارد [۲۳]. لی در سال ۲۰۰۶ از الگوریتم سیفت برای ثبت تصاویر که تغییرات روشنایی داشتند استفاده کرد و به طور موفقیت‌آمیزی تعداد زیادی عوارض نقطه‌ای پایدار را استخراج کرد [۲۴]. در مقاله حاضر نیز از الگوریتم سیفت برای استخراج عوارض استفاده شده است.

#### ۳-۴-۳- تناظریابی

در این مرحله مقایسه یک بردار توصیفگر از تصویر مرجع با تمام بردارهای توصیفگر نقاط کلیدی در تصویر هدف، به صورت ضرب داخلی دو بردار انجام می‌پذیرد. مفهوم ضرب داخلی در اینجا بر اساس آنالیز برداری، همان مفهوم حداقل فاصله اقلیدسی را دارد. از نتیجه یک ضرب داخلی بردارهای توصیفگر دو تصویر، یک ماتریس برداری  $n \times 1$  حاصل می‌شود که n تعداد نقاط کلیدی در تصویر هدف می‌باشد و سپس آن را از کمترین به بیشترین مقدار مرتب می‌کنیم. برای هر بردار توصیفگر در تصویر هدف، اولین و دومین توصیفگر نزدیک در توصیفگرهای تصویر مرجع شناخته می‌شوند. یک زوج توصیفگر نزدیک، یک زوج نقطه کلیدی متناظر را مشخص می‌کند. اگر نسبت فواصل اولین عنصر کمترین به دومین عنصر کمترین در بردار ضرب داخلی از یک آستانه‌ای کمتر باشد آن دو نقطه کلیدی در دو تصویر را به عنوان دو نقطه متناظر انتخاب می‌کنیم [۲۵ و ۲۶].

#### ۳-۴-۳- حذف نقاط اشتباہ

به علت اینکه بعضی از انطباق‌های اشتباہ در مرحله تناظریابی تصاویر ممکن است رخ دهد، با استفاده از

<sup>1</sup> Ransac

<sup>2</sup> Fischler , Bolles

<sup>3</sup> Inlier

<sup>4</sup> Outlier

به هنگام سازی انتخابی است. در میانگین‌گیری، پس زمینه به شیوه‌ای تدریجی<sup>۲</sup> با میانگین‌گیری از پس زمینه قبلی و فریم جاری، ساخته می‌شود. اگر یک میانگین وزن دار را بر پس زمینه قبلی و فریم جاری اعمال کنیم، تصویر پس زمینه به روش به هنگام سازی انتخابی ساخته شده است.

روش دیگر استخراج تصویر پس زمینه استفاده از فیلتر میانه<sup>۳</sup> است. لازم به ذکر است که هریک از روش‌های ذکر شده دارای مزايا و معایبی می‌باشند. به عنوان مثال از آنجاکه میانگین شدت رنگ مجموعه‌ای از پیکسل‌های تشکیل دهنده یک تصویر ممکن است عددی خارج از مجموعه باشد، درنتیجه این فیلتر باعث از بین رفتن جزئیات موجود در تصاویر خواهد شد. اما از آنجا که میانه شدت رنگ مجموعه‌ای از پیکسل‌های تشکیل دهنده یک تصویر، عددی در داخل همان مجموعه می‌باشد، درنتیجه فیلتر میانه جزئیات موجود در تصاویر را حفظ می‌کند. درنتیجه در این مقاله از روش فیلتر میانه برای استخراج تصویر پس زمینه استفاده شده است.

پس از استخراج تصویر پس زمینه تمام فریم‌های ورودی مطابق روش شکل (۱) با تصویر پس زمینه ثبت می‌شوند. با این کار تلاش می‌شود که مشکل حرکات اضافی دوربین تا حد ممکن کاهش یافته و دقت تشخیص خودروها افزایش یابد.

بعدازین که عملیات ثبت روی فریم‌ها صورت گرفت برای شناسایی عوارض متحرک، هر تصویر ورودی با تصویر پس زمینه تفاصل مطابق رابطه (۲) گرفته می‌شود.

رابطه (۲)  $Dif_n = |updated\ background_p - F_n|$  <sup>*updated background<sub>p</sub>*</sup> تصویر پس زمینه به روزرسانی مرحله p ام، و  $F_n$  تصویر هم مرجع شده n ام می‌باشد.

مقیاس و دوران و عمود نبودن محورها می‌باشد.]

#### ۴- تشخیص خودروها

برای تشخیص خودروها به این صورت عمل می‌شود که بعد از اینکه فریم‌ها استخراج شدن، از بین این تصاویر، تصویر پس زمینه<sup>۱</sup> محاسبه خواهد شد و کل تصاویر ورودی از پس زمینه کم می‌کنیم تا عوارض متحرک از تصویر استخراج شوند. زمانی که دوربین دارای حرکات اضافی باشد الگوریتم‌هایی که از مدل سازی پس زمینه یا تفاصل فریم برای به دست آوردن ناحیه‌های مرتبط به اشیاء متحرک استفاده می‌کنند ناحیه‌های نادرستی (عوارض ثابت) را از پس زمینه به عنوان اشیاء متحرک برچسب گذاری می‌کنند. همچنین ناحیه‌های نادرستی را از پس زمینه به عنوان اشیاء متحرک برچسب گذاری می‌کنند که در واقع ثابت هستند و حرکت دوربین باعث ایجاد خطأ در این موارد می‌شود. برای جبران این حرکات اضافی دوربین بعد از استخراج پس زمینه تمام فریم‌های جاری را نسبت به پس زمینه ثبت می‌کنیم و بعد با استفاده از روش تفاصل فریم‌ها به شناسایی خودروها پرداخته می‌شود. در شکل (۲) فلوچارت پیشنهادی برای شناسایی خودروها نشان داده شده است.

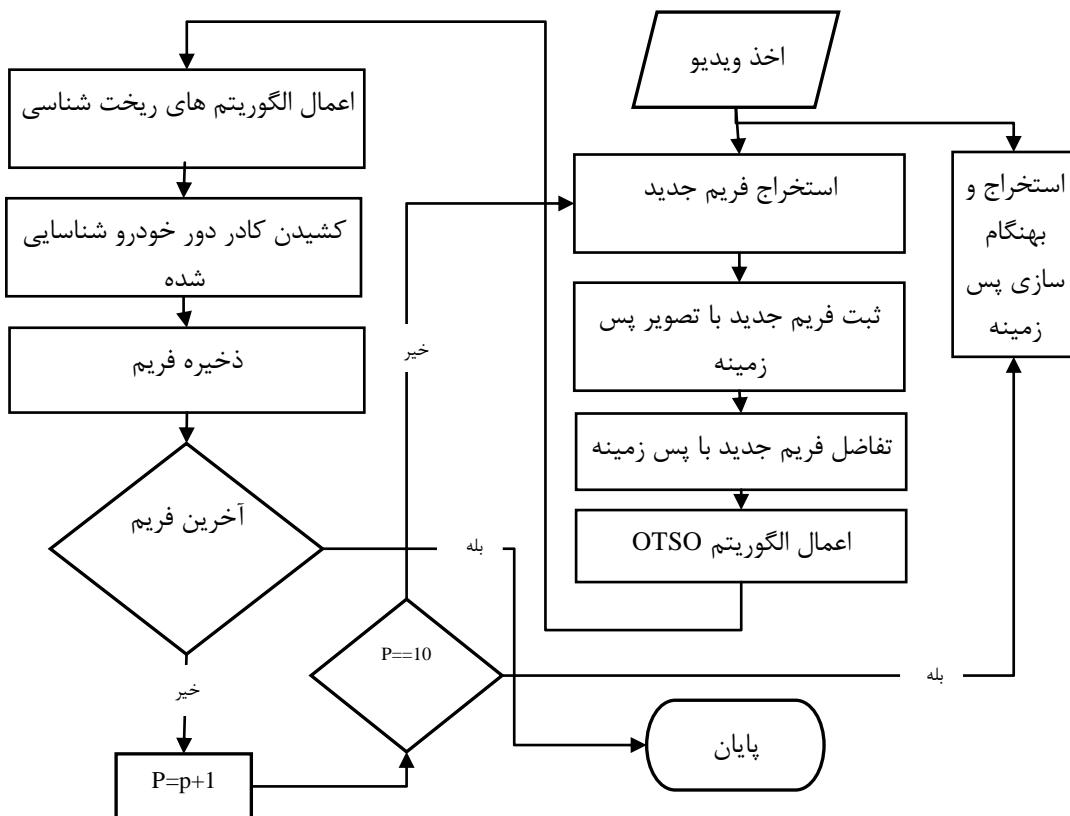
برای استخراج خودروها، همان طور که در شکل (۲) نشان داده شده است، بعد از اینکه تصاویر از فیلم ویدیویی استخراج شد اگر یک پس زمینه ثابت نداشته باشیم که عاری از خودروها و عوارض متحرک دیگر باشد باید یک تصویر پس زمینه از فریم‌های ویدیو استخراج کرد.

پس زمینه ممکن است به میزان زیادی توسط واقع شدن سایه ساختمان‌ها یا ابرها و یا به دلیل تغییر در شرایط روشنایی صحنه، تغییر یابد. با این شرایط محیطی متغیر، تصویر پس زمینه باید به طور منظم به هنگام شود. چندین روش برای به هنگام سازی تصویر پس زمینه وجود دارد. متدائل ترین آن‌ها میانگین‌گیری و

<sup>2</sup> Gradual

<sup>3</sup> Median Filter

<sup>1</sup> Background



شکل ۲: الگوریتم پیشنهادی برای شناسایی خودروها

روش ارائه شده توسط اتسو<sup>2</sup> که یکی از روش‌های معروف در آستانه‌گیری سراسری است<sup>3</sup> و بر اساس هیستوگرام تصویر عمل می‌نماید، کاربرد فراوانی در پردازش تصویر دارد. در این پژوهش از روش اتسو استفاده شده است [۲۹].

پس از آستانه‌گیری اتسو، عوارض متحرک که شامل نویز هم می‌باشند، در تصویر دودویی باقی می‌مانند. شکل خودروها کاملاً یکپارچه نبوده و در بعضی نقاط قسمت‌های مختلف یک خودرو در حال جدا شدن می‌باشند. برای رفع این نواقص از عملگرهای ریخت‌شناسی استفاده می‌کنیم.

اولین عملگر ریخت‌شناسی که استفاده خواهیم کرد، عملگر بستن می‌باشد. این عملگر باعث کوچک‌تر شدن و جمع شدن اشیای موجود در تصویر می‌شود.

در تصویر تفاضل ممکن است نواقصی در خودروهای شناسایی شده وجود داشته باشد، به عنوان مثال ممکن است دو خودرو به یکدیگر متصل شده باشند و یا قسمت‌های مختلف یک خودرو از یکدیگر جدا شده و منجر به ایجاد خطاهایی در مراحل بعدی پردازش به خصوص در مرحله تشخیص خودروها شوند. بنابراین در ادامه کار از روش‌های ریخت‌شناسی<sup>1</sup> موجود در حوزه پردازش تصویر جهت برطرف کردن این نواقص بهره خواهیم گرفت. مراحل به کار گرفته شده برای بهبود گام تشخیص خودروها به ترتیب در ادامه شرح داده خواهند شد.

ابتدا جهت ساده شدن پردازش‌های بعدی، آستانه‌ای بر روی تصویر تفاضلی برای تبدیل این تصویر سطح خاکستری به تصویری دوسری اعمال خواهد شد.

<sup>2</sup> Otsu<sup>3</sup> Global thresholding<sup>1</sup> Morphology

خودرو استخراج می‌شود. با استفاده از مستطیل‌های محاطی، دور خودروهای شناسایی شده در فریم اصلی یک پنجره کشیده می‌شود.

### ۵- پیاده‌سازی

برای پیاده‌سازی الگوریتم پیشنهادی از نرم‌افزار متلب استفاده شده است. داده‌های مورداستفاده برای تست روش پیشنهادی، از سه پایگاه داده گرفته شده‌اند که مشخصات مربوط به هر کدام از این داده‌ها در جدول (۱) تشریح شده است.

همچنین این عملگر منجر به پرسدن فضای خالی ایجاد شده در میان اشیای استخراج شده می‌شود. پس از اعمال عملگر بستن، مؤلفه‌های همبند را از درون تصویر استخراج می‌کنیم. درواقع مؤلفه‌های همبند، نمایشی از ناحیه اشغال شده توسط خودروها در تصویر می‌باشند که این ناحیه یک همسایگی را در فضای تصویر تعریف می‌کند. سپس برای به دست آوردن مراکز تقریبی خودروها، بر روی هریک از خودروهای استخراج شده مستطیلی محاطی رسم می‌کنیم و سپس مرکز این مستطیل به عنوان مرکز نقل یا موقعیت

جدول ۱: داده‌های مورد استفاده برای ارزیابی سیستم پیشنهادی

| فرمت | تعداد فریم بر ثانیه | اندازه فریم (پیکسل) | مدت زمان فیلم (دقیقه) | تراکم خودروها | میزان لرزش دوربین | مکان                           | نام داده |
|------|---------------------|---------------------|-----------------------|---------------|-------------------|--------------------------------|----------|
| MP4  | ۲۵                  | ۳۶۰*۴۸۰             | ۲۰                    | متوسط         | متوسط             | یکی از بزرگراه‌های کشور اکراین | Highway1 |
| MOV  | ۲۵                  | ۷۲۰*۱۲۸۰            | ۲                     | زیاد          | زیاد              | بزرگراه رسالت در شهر تهران     | MVI_01   |
| MOV  | ۲۵                  | ۷۲۰*۱۲۸۰            | ۲                     | کم            | متوسط             | بزرگراه رسالت در شهر تهران     | MVI_02   |

جمعه‌ای‌بار پردازش ویدئویی نرم‌افزار متلب، به فریم‌های متوالی تبدیل کردیم تا پردازش‌های بعدی را بر روی این تصاویر استخراج شده انجام دهیم. شکل (۳) نمونه‌ای از فریم‌های استخراج شده را نشان می‌دهد.

ویژگی اصلی این پایگاه‌های داده آن است که دوربین نظارتی برای باد و فشارهای جانبی دچار لرزش‌های متوالی شده که در تشخیص خودروها می‌تواند مشکل ایجاد کند. ویدئوهای موردنظر را با استفاده از امکانات ارائه شده در



شکل ۳: نمونه‌ای از فریم‌های استخراج شده

پس زمینه استخراج شده توسط این روش برای سه ویدیویی مورد بررسی را نشان می‌دهد.

برای استخراج تصویر پس زمینه در این تحقیق از فیلتر میانه به دلیل حفظ جزئیات موجود در تصویر استفاده شده است. شکل (۴) نمونه‌ای از تصاویر



شکل ۴: نمونه‌ای از تصاویر پس زمینه استخراج شده برای سه ویدیو

سیفت تناظریابی صورت گرفته و بردارهای ویژگی به دست می‌آیند. از طریق این بردار ویژگی و همچنین یکتابع تبدیل افاین می‌توان این تصاویر را بر روی هم ثبت کرد. در شکل (۵) نتیجه ثبت تصاویر با پس زمینه نشان داده شده است.

#### ۵-۱- جبران حرکت از طریق ثبت تصاویر به پس زمینه

پس از استخراج تصویر پس زمینه تمام فریم‌های ورودی را یک به یک با تصویر پس زمینه ثبت می‌کنیم. برای تصویر پس زمینه و هر فریم ورودی با الگوریتم



شکل ۵: سطر اول) فریم ورودی. سطر دوم) میزان تغییرات فریم ورودی نسبت به پس زمینه. سطر سوم) فریم ثبت شده با تصویر پس زمینه

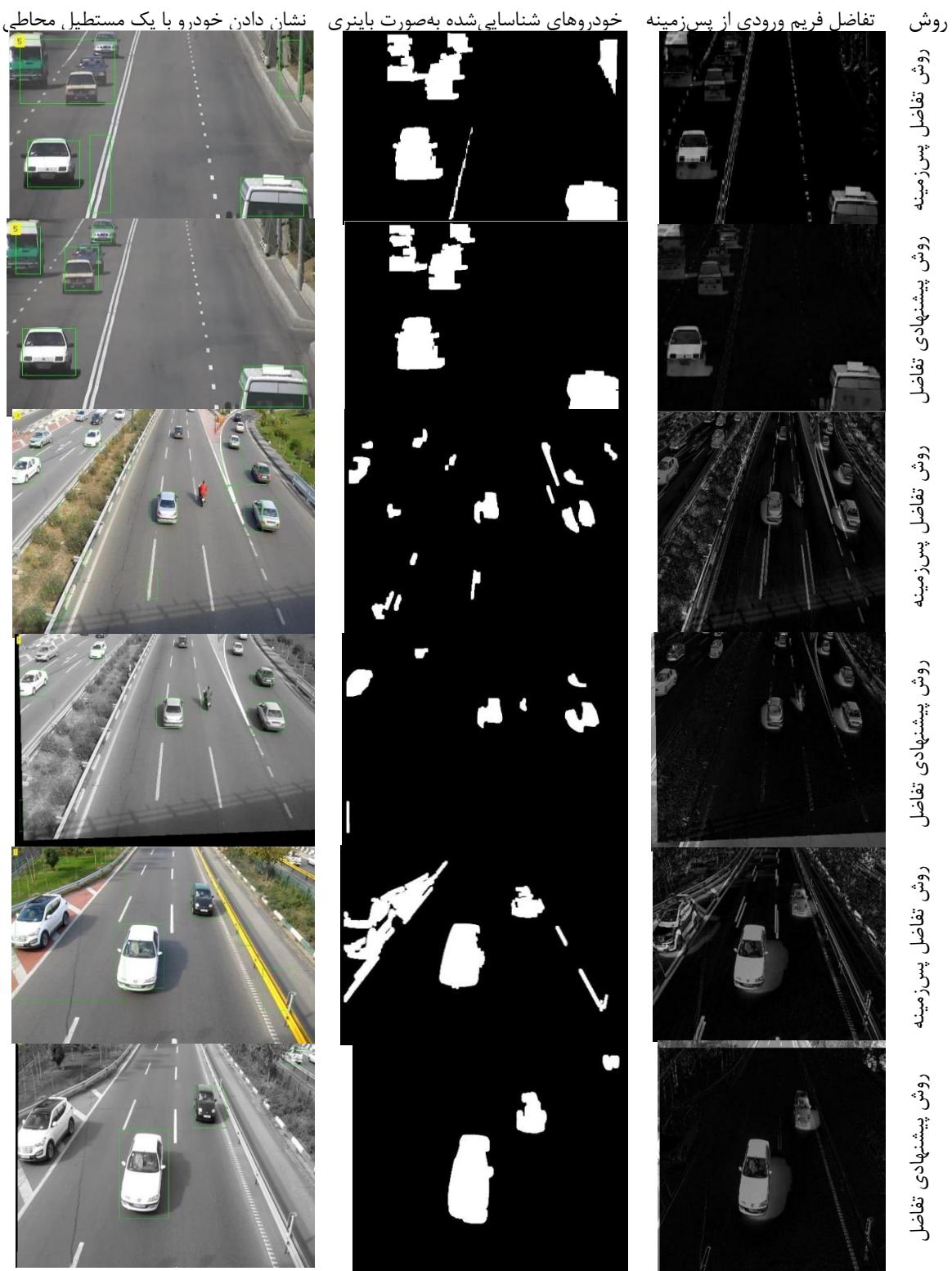
## ۲-۵- استخراج خودروها

در این تحقیق برای استخراج خودروها در فریم‌های متوالی، از روش تفاضل پس زمینه استفاده شده است. بدین ترتیب که پس از استخراج تصویر پس زمینه، این تصویر از هریک از فریم‌های ورودی تفريقي خواهد شد و حاصل این تفريقي تصویری خواهد بود که در آن تنها عوارض غیر ثابت (خودروها و ...) حضور دارند و درنتیجه خودروها به همراه برخی عوارض ثابتی که در اثر حرکات اضافی دوربین به اشتباه به عنوان عوارض متحرک شناسایی می‌شوند، استخراج خواهند شد.

در شکل (۵) ستون‌ها داده‌های مورد استفاده را نشان می‌دهد و سطرها نتایج مرحله به مرحله الگوریتم پیشنهادی برای جبران حرکت دوربین را نشان می‌دهد. سطر اول تصاویر اصلی مربوط به هر سه نوع داده یاد شده در جدول (۱) را نشان می‌دهد. سطر دوم سطر دوم میزان جابجایی فریم جاری نسبت به پس زمینه را نشان می‌دهد. در این اشکال همانطور که نشان داده شده است، پیکسل‌های که نسبت به پس زمینه دارای تغییر می‌باشد به صورت رنگی نشان داده شده است. سطر سوم نتایج ثبت تصاویر ورودی با پس زمینه را نشان می‌دهد.

## جبران حرکت دوربین‌های کنترل نظارتی با استفاده از...

علی کرمی، محسن سریانی، مسعود ورشوساز



شکل ۶: ستون اول) تفاضل فریمهای متوالی از تصویر پس زمینه در هر دو حالت با ثبت و بدون ثبت. ستون دوم) خودروهای شناسایی شده به صورت باینری در هر دو حالت با ثبت و بدون ثبت. ستون سوم) نشان دادن خودرو با یک مستطیل محاطی در هر دو حالت با ثبت و بدون ثبت

دوربین تا حد خوبی جبران شده است (شکل(۶):  
سطرهای زوج).

#### ۶- ارزیابی

نتایج هیچ پژوهشی قابل اعتماد نیست، مگر زمانی که صحتش ارزیابی شده باشد. صحت، یعنی سطح تطابق بین برچسب‌های پیکسل‌های خروجی از روش پیشنهادی و برچسب‌های مبتنی بر داده‌های مرجع، که با نام داده صحیح مرجع شناخته می‌شود [۳۰، ۳۱].

معیارهای مناسب برای صحت استخراج خودرو را می‌توان با معیارهای عملکرد طبقه‌بندها فراهم ساخت چون تصاویر خروجی شناسایی خودرو درواقع به دو کلاس صفر و یک تقسیم‌بندی می‌شوند که کلاس یک مربوط به خودرو و کلاس صفر مربوط به غیر خودرو می‌شود ابزار معمولی که برای ارزیابی صحت طبقه‌بندی استفاده می‌شود بر حسب یک ماتریس ابهام (یا خطای می‌باشد. یک ماتریس ابهام، یک آرایه مرربع با ابعاد  $n \times n$  می‌باشد، که تعداد کلاس‌ها را نشان می‌دهد و در این تحقیق فقط دو کلاس داریم. این ماتریس، رابطه بین دو نمونه از اندازه‌گیری‌ها از منطقه‌ای که برچسب‌دهی شده است را نشان می‌دهد. مجموعه اول داده مرجع را نشان می‌دهد که به صورت دستی خودروها استخراج شده‌اند. مجموعه دوم، از برچسب پیکسل‌هایی تشکیل شده است که به وسیله روش پیشنهادی استخراج شده‌اند جدول (۲).

در شکل (۶) نمونه‌ای از تفاضل پس‌زمینه با فریم‌های متوالی برای استخراج اشیای متحرک موجود در تصویر نشان داده شده است. ستون اول تفاضل فریم‌های متوالی از تصویر پس‌زمینه در هر دو حالت با ثبت و بدون ثبت را نشان می‌دهد. در ستون دوم خودروهای شناسایی شده به صورت باینری در هر دو حالت با ثبت و بدون ثبت دیده می‌شوند. در ستون سوم خودروهای شناسایی شده با یک مستطیل محاطی در هر دو حالت با ثبت و بدون ثبت نشان داده شده‌اند.

جهت ساده شدن پردازش‌های بعدی، آستانه‌ای بر روی تصویر تفاضل برای تبدیل این تصویر سطح خاکستری به تصویری دوستخی، اعمال خواهد شد. برای اعمال این آستانه از روش رایج اتسو که آستانه بینه را پیدا می‌کند استفاده شده است. برای حذف نواحی کوچک حاصل از نویزها از الگوریتم‌های ریخت شناسی<sup>۱</sup> مانند بازکردن و بستن با ماسک ۹\*۹ استفاده شده است(شکل(۶) ستون دوم). پس برای به دست آوردن دقیق‌تر مراکز خودروها، بر روی هریک از خودروهای استخراج شده به صورت اتوماتیک مستطیلی محاطی رسم شده و سپس مرکز این مستطیل محاطی به عنوان مرکز نقل یا موقعیت خودرو استخراج می‌شود(شکل(۶) ستون سوم).

همانطور که در سطرهای فرد شکل(۶) پیداست به دلیل حرکات اضافی که دوربین در حین فیلمبرداری داشته است، دو خطای در تشخیص صحیح عوارض متحرک ایجاد شده است. اول اینکه زمانی که خودروها به هم نزدیک باشند، حرکت اضافی دوربین باعث می‌شود که در تصویر تفاضلی خودروهای نزدیک باهم تلفیق شده و به عنوان یک خودرو شناسایی شوند. دوم اینکه ناحیه‌های نادرستی(عوارض ثابت) از پس‌زمینه به عنوان اشیاء متحرک برچسب گذاری می‌شوند. ولی با اعمال ثبت تصاویر حرکات اضافی

<sup>1</sup> Morphology

جدول ۲: ماتریس ابهام

| نقشه مرجع (واقعیت مرجع) |              |             | ماتریس ابهام |                            |
|-------------------------|--------------|-------------|--------------|----------------------------|
| مجموع                   | شناسایی نشده | شناسایی شده | شناسایی شده  | نتایج حاصل از روش پیشنهادی |
| TP+FP                   | (FP)         | (TP)        | شناسایی نشده | TP                         |
| TN+FN                   | (TN)         | (FN)        | شناسایی نشده | TN                         |
| FP+TP+TN+FN             | FP+TN        | TP+FN       | مجموع        | FP                         |

تشخیص درست به وسیله تقسیم درایه (j.i) به مجموع ستون  $\alpha$  محاسبه می‌شود. نرخ تشخیص درست رابطه (۴)، در حقیقت، معیاری از شاخص جامع بودن<sup>۷</sup> (کامل بودن) است که نشان دهنده‌ی کسری از پیکسل‌های واقعی کلاس مورد نظر (کلاس مربوط به خودروها) است که در روش پیشنهادی به درستی در کلاس خود قرار گرفتند. به بیان دیگر کامل بودن، بیانگر درصدی از پیکسل‌های کلاس مرجع است که به درستی در کلاس استخراج شده وجود دارد.

$$HR = \frac{TP}{TP+FN}$$

رابطه (۴)  
 ۳) نرخ تشخیص اشتباه<sup>۸</sup>

نرخ تشخیص اشتباه رابطه (۵) تقسیم مجموع تعداد پیکسل‌های شناسایی شده توسط الگوریتم پیشنهادی که به اشتباهی طبقه‌بندی شده است بر کل پیکسل‌ها را نشان می‌دهد [۳۰].

$$FAR = \frac{FP+FN}{TP+FN+TN+FP}$$

رابطه (۵)  
 ۴) معیار دقت کشف چند شیئی MODP<sup>۹</sup> معیار دیگری برای ارزیابی روش پیشنهادی در شناسایی خودرو MODP می‌باشد [۳۲]. در این معیار از رابطه همپوشانی خاصی بین داده‌های مرجع و روش پیشنهادی، مطابق رابطه (۶) استفاده شده است.

$$Overlap\ Ratio = \sum_{i=1}^{N_{mapped}^t} \left| \frac{G_i^{(t)} \cap D_i^{(t)}}{G_i^{(t)} \cup D_i^{(t)}} \right|$$

$$MODP = \frac{Overlap\ Ratio}{N_{mapped}^t}$$

در این رابطه،  $G_i^{(t)}$  و  $D_i^{(t)}$  بیانگر عارضه واقعی و

ستون‌های ماتریس ابهام داده مرجع را نشان می‌دهد، و سطرهای آن، برچسب‌های اختصاص داده شده توسط روش پیشنهادی را نشان میدهد. در این ماتریس<sup>۱</sup> TP تعداد نقاطی است که در هر دو مجموعه داده‌ها برچسب توصیفی برای آن شناسایی شده است، در حالی که TN<sup>۲</sup> تعداد نقاطی را مشخص می‌کند که در هر دو داده مورد شناسایی واقع نشده باشد. همچنین FP<sup>۳</sup> تعداد نقاطی هستند که توسط روش پیشنهادی به اشتباه شناسایی شده‌اند. در نهایت FN<sup>۴</sup> تعداد نقاطی است که توسط روش پیشنهادی به اشتباه مورد شناسایی واقع نشده است.

۱) شاخص دقت کلی<sup>۵</sup>

می‌توان چندین شاخص برای صحت شناسایی خودرو را از ماتریس ابهام به دست آورد. مطابق رابطه (۳) صحت کل (OA) از نسبت تعداد پیکسل‌های خودروهایی که به درستی به کلاس خودرو تعلق گرفته‌اند به تعداد کل پیکسل‌ها یا خودروها محاسبه می‌شود [۳۱, ۳۰].

رابطه (۳)

$$OA = (TP + TN) / (TP + FN + TN + FP)$$

در حقیقت رابطه فوق برابر است با تقسیم مجموع درایه‌های قطر اصلی ماتریس ابهام به تعداد کل نمونه‌ها.

۲) نرخ تشخیص درست<sup>۶</sup>

برای هر کلاس اطلاعاتی در ماتریس ابهام، نرخ

<sup>1</sup> True Positive

<sup>2</sup> True Negative

<sup>3</sup> False Positive

<sup>4</sup> False Negative

<sup>5</sup> Overall Accuracy

<sup>6</sup> Hit Rate

<sup>7</sup> Completeness

<sup>8</sup> False Alarm Rate

<sup>9</sup> Multi Object Detection Precision

پس زمینه تمام فریم‌های متوالی را از پس زمینه کم کرده و عوارض متحرک استخراج شده است. حالت دوم بعداز اینکه تصویر پس زمینه استخراج شد، تمام فریم‌های متوالی با تصویر پس زمینه ثبت شده و بعد از تفاضل گرفته می‌شود.

نتایج به دست آمده در هر دو حالت قبل از ثبت و بعد از ثبت در جدول (۳) نشان داده شده است.

استخراج شده آم در فریم آم می‌باشد. همچنین  $N_{mapped}^t$  مشخص کننده تعداد عارضه تصویر شده در فریم آم می‌باشد.

برای ارزیابی توانایی الگوریتم پیشنهادی در دو حالت مختلف روی داده‌های موجود در جدول (۱) به استخراج خودروها پرداخته شده است. برای این کار در حالت اول از روی فریم‌های ورودی یک تصویر پس زمینه استخراج می‌شود. سپس با استفاده از روش تفاضل

جدول ۳: نتایج حاصل شده برای شناسایی خودرو در دو حالت مختلف با ثبت و بدون ثبت

| نام داده                               | تعداد فریم | تعداد خودروها | تعداد کل | تشخیص صحیح | تشخیص غلط(تلفیق و بر عکس) | غیرقابل تشخیص | تثبیت عارضه به عنوان خودرو | دقت کلی (درصد) | روش کار  |
|--|------------|---------------|----------|------------|---------------------------|---------------|----------------------------|----------------|----------|
| ۱MVI_000                               | ۵۰۰        | ۴۴۳۰          | ۳۹۵۰     | ۴۶۵        | ۷۵۷                       | ۳۱            | ۴۲۵                        | ۶۸,۸۹          | بدون ثبت |
|  | ۵۰۰        | ۸۵۲           | ۳۳۴۵     | ۶۲         | ۷۸۹                       | ۱             | ۵۲                         | ۸۸,۱۳          | با ثبت   |
| ۲MVI_000                               | ۵۰۰        | ۱۷۲۷          | ۷۶۲      | ۸۸         | ۱۷۶                       | ۵             | ۲۸                         | ۸۸,۰۹          | با ثبت   |
|  | ۱۵۰۰       | ۷۰۰۹          | ۱۵۴۶     | ۲۵۶        | ۱۴۶۶                      | ۵             | ۱۴۲                        | ۷۸,۴۴          | بدون ثبت |
| Highway1                               | ۵۰۰        | ۶۲۸۵          | ۲۹۸      | ۱۱۰۱       | ۵۵۷۳                      | ۳۸            | ۶۴۷                        | ۷۶,۳۶          | با ثبت   |
|  | ۱۵۰۰       | ۷۰۰۹          | ۲۱۲      | ۵۶۱        | ۲۹۸                       | ۱۷            | ۴۲۵                        | ۸۸,۱۳          | بدون ثبت |
| قدرت مطلق تفاضل حالت با ثبت و بدون ثبت |            |               |          |            |                           |               |                            |                |          |

اولاً تعداد خودروهای صحیح تشخیص داده شده قبل از اعمال الگوریتم پیشنهادی ۵۵۷۳ و بعد از اعمال الگوریتم پیشنهادی ۶۲۸۵ بوده است. که باعث تشخیص ۲۱۲ خودرو بیشتر از حالت قبل است.

در روشنی که از ثبت تصویر استفاده نمی‌شود هنگامی که چندین خودرو کنار هم یا نزدیک به هم قرار گرفته‌اند در اثر لرزش‌هایی که به دوربین وارد شده باعث شده که خودروهای نزدیک به هم با هم تلفیق شوند و چند خودرو را به عنوان یک خودرو شناسایی کند و هر چقدر تراکم خودروها زیاد باشد با

مطابق جدول (۳) تعداد کل خودروهای موجود در تمام فریم‌ها، تعداد خودروهای شناسایی شده صحیح در دو حالت، تعداد خودروهایی که به اشتباہ ادغام و یا تجزیه شده‌اند، تعداد خودروهای غیرقابل تشخیص و در نهایت تعداد عارضه ثابتی که به اشتباہ به عنوان عارضه متحرک شناسایی شده، نشان داده شده است. براساس این اطلاعات برای هر کدام از داده‌ها مطابق رابطه (۳) دقیق کلی در شناسایی خودروها محاسبه شده است. با مقایسه‌ای که بین دو حالت مختلف (با ثبت و بدون ثبت تصاویر) صورت گرفته می‌توان نتیجه گرفت که

MVI\_01، هم میزان خودروهایی که تلفیق شده‌اند بیشتر بوده و هم اینکه تعداد عوارض ثابتی که به اشتباه به عنوان عارضه متحرک تشخیص داده شده‌اند، MODP زیاد می‌باشد. در نتیجه تفاوت نمودارهای HR، FAR و QAR قبل از ثبت و بعد از ثبت بر جسته می‌باشد. این نتایج نشان می‌دهد که هر چه حرکات اضافی دوربین حین فیلمبرداری بیشتر باشد و همچنین تراکم خودروها بیشتر باشد، دقت و قابلیت اعتماد در شناسایی خودروها با استفاده از روش تفاضل فریم‌ها کمتر می‌شود ولی اگر قبل از عمل تفاضل تمام فریم‌ها را به پس‌زمینه ثبت کرد و بعد به استخراج خودروها پرداخت، در آن صورت تا حد خوبی دقت و قابلیت اعتماد در شناسایی خودروها بالا می‌رود.

در مجموع و به طور کلی دقت نهایی (OA) تشخیص خودروها در حالت اول یعنی قبل از ثبت ۷۶٪ و با روش پیشنهادی یعنی بعد از ثبت ۸۹٪ می‌باشد. این نتایج نشان می‌دهد که عمل ثبت تصاویر متواالی با پس‌زمینه حدود ۱۳٪ در تشخیص خودروها بهتر عمل کرده است.

## ۷-نتیجه‌گیری

در این تحقیق روشی مناسب برای جبران حرکت اضافی دوربین‌های نظارتی در اثر عوامل بیرونی نظری باد و غیره با استفاده از ثبت تصاویر متواالی با پس‌زمینه در شناسایی خودروها، ارائه شده است. برای این‌منظور بعد از استخراج فریم‌ها و پیدا کردن تصویر پس‌زمینه با استفاده از روش میانه‌گیری، باید تمام فریم‌های ورودی با تصویر پس‌زمینه بر اساس الگوریتم پیشنهادی که در بخش (۳) ارائه شده، ثبت شوند. سپس با استفاده از روش تفاضل پس‌زمینه به شناسایی خودروها پرداخته شده است. این امر باعث حذف خطای دیلای در حین تصویربرداری می‌شود. برای ارزیابی نتایج، روش پیشنهادی بر روی سه نوع داده پیاده‌سازی شده است. ویژگی اصلی این مجموعه داده‌ها لرژش‌های مکرر دوربین در حین فیلمبرداری و

لرژش دوربین عمل تلفیق خودروها بیشتر اتفاق می‌افتد. علاوه بر اینکه خودروها را با هم تلفیق می‌کند باعث می‌شود عوارض ثابتی که درجه خاکستری آن‌ها با محیط زیاد فرق دارند (مثل خط‌کشی خیابان‌های شهری و ...) در دو فریم متواالی به عنوان عوارض متحرک شناسایی شوند. در مجموع تعداد عوارض ثابتی که به عنوان عوارض متحرک شناسایی شده، در حالت اول ۶۴۷ عارضه می‌باشد که بعد از ثبت تصاویر این تعداد به ۸۶ عارضه کاهش یافته است که این خود نشان دهنده توانایی ثبت تصاویر متواالی با تصویر پس‌زمینه در موقعی که دوربین نظارتی دچار لرژش‌های مکرر شده را نشان می‌دهد.

در ادامه برای بدست آوردن میزان دقت و قابلیت اعتماد در تشخیص خودروها از معیارهای دیگری که در روابط (۳، ۴، ۵ و ۶) بیان شده استفاده می‌شود. برای این کار به صورت اتفاقی ۱۰ فریم از هر مجموعه داده را انتخاب می‌کیم. در دو حالت بدون ثبت و با ثبت تصاویر، خودروها استخراج می‌شوند. پیکسل‌های مربوط به موقعیت دقیق خودروها در این ۱۰ فریم به صورت دستی استخراج شده و به عنوان داده‌های واقعی در ارزیابی مورد استفاده قرار می‌گیرند. بعد از این مرحله به تهیه ماتریس ابهام برای هر ۱۰ فریم در هر سه داده و در دو حالت مختلف با ثبت و بدون ثبت پرداخته می‌شود.

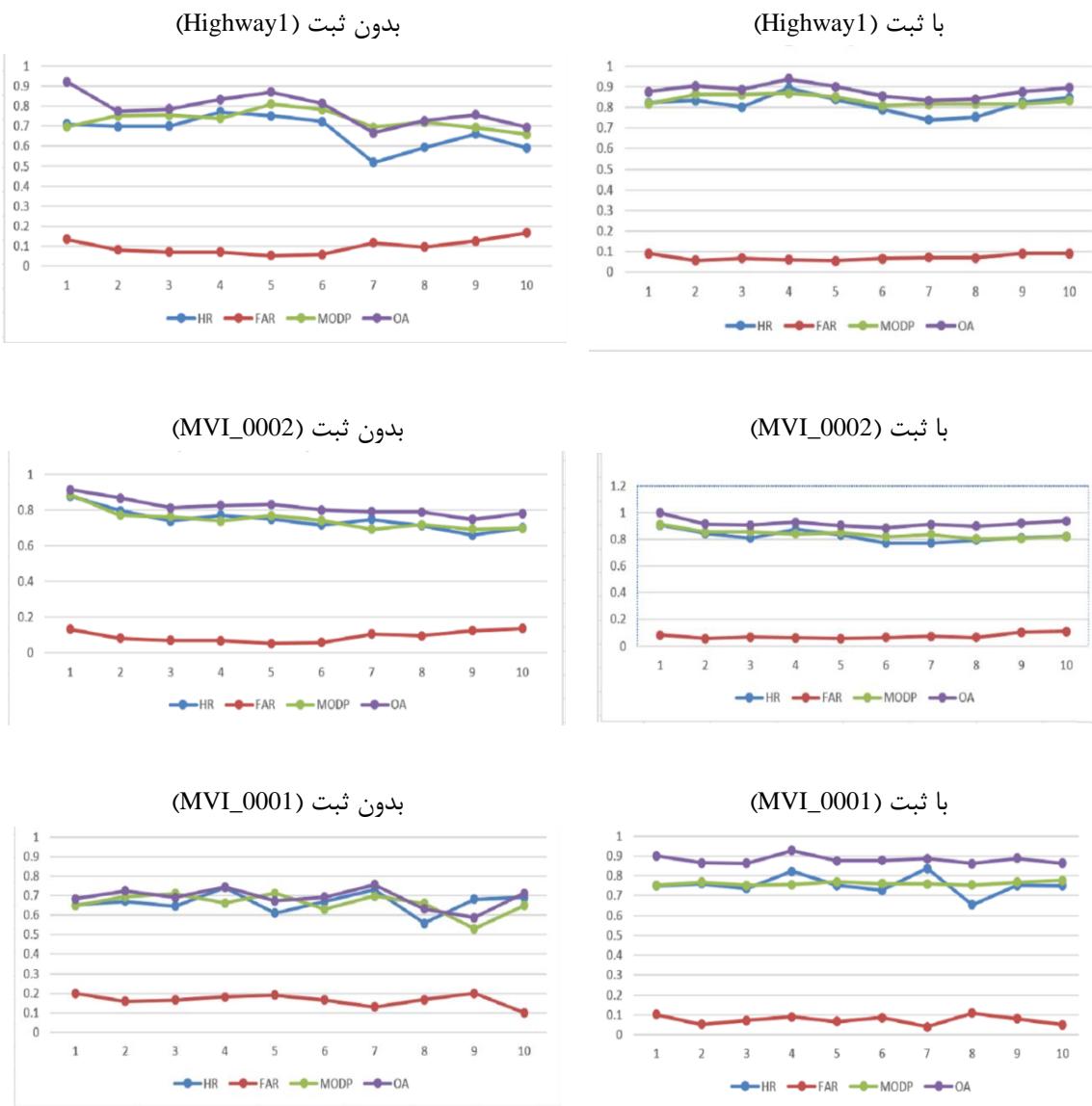
بطور کلی برای ارزیابی دقت در این تحقیق از ۶۰ ماتریس ابهام ( $2 \times 3 \times 10$ ) استفاده شده است. نتایج در قالب نمودارهای شکل (۷) نشان داده شده است.

بطور کلی با توجه به نمودارهای شکل (۷) میزان HR و MODP بعد از ثبت بیشتر از حالتی است که عمل ثبت صورت نگرفته است و همچنین میزان FAR هم بعد از ثبت نسبت به قبل از ثبت کمتر شده است و این خود نشان دهنده توانایی الگوریتم پیشنهادی در شناسایی دقیق‌تر خودروها می‌باشد.

همان‌طور که در جدول (۳) نشان داده شده است به دلیل تراکم بالای خودروهای موجود در داده‌های

این تعداد به ۸۶ عارضه کاهش پیداکرده است و همچنین میزان HR و MODP بعد از ثبت، بیشتر و FAR هم کمتر شده است که این نشان بر توانایی روش پیشنهادی در شناسایی خودروها می‌باشد. نتایج به دست آمده در این تحقیق نشان داد که دقت کلی روش پیشنهادی نسبت به حالت اول (بدون عمل ثبت) حدود ۱۳ درصد در شناسایی عوارض متحرک بیشتر شده است.

متفاوت بودن تراکم خودروهای موجود در هر مجموعه داده می‌باشد. در دو حالت بدون استفاده از ثبت و با استفاده از ثبت تصویر به شناسایی خودروها پرداخته شده است. در هر دو حالت با استفاده از متریک‌های OA, MODP, HR, FAR و OA، مقایسه صورت گرفته است. در حالت اول تعداد عوارض ثابتی که به عنوان خودرو شناخته شده است حدود ۶۴۷ عارضه می‌باشد که با اعمال الگوریتم پیشنهادی



شکل ۷: نتایج به دست آمده برای ۱۰ فریم از ۳ مجموعه تصاویر ویدیویی برای هر دو حالت با و بدون ثبت تصاویر

کارآمدترین روش برای جبران حرکت اضافی دوربین با استفاده از روش تفاضل فریم با پس زمینه پرداخته شود.

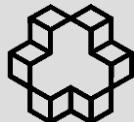
## ۸- پیشنهادات

به منظور ادامه و تکمیل تحقیقات صورت گرفته در زمینه جبران حرکت اضافی دوربین پیشنهاد می‌گردد که با مقایسه روش‌های ثبت تصاویر به معروفی بهترین و

## مراجع

- [1] A. Karamiani and N. Farajzadeh, "Optimal feature points for tracking multiple moving objects in active camera model," *Multimedia Tools and Applications*, pp. 1-19, 2015.
- [2] J. K. Aggarwal and Q. Cai, "Human motion analysis: A review," in *Nonrigid and Articulated Motion Workshop*, 1997. Proceedings., IEEE, 1997, pp. 90-102.
- [3] M. Piccardi, "Background subtraction techniques: a review," in *Systems, man and cybernetics, 2004 IEEE international conference on*, 2004, pp. 3099-3104.
- [4] R. Poppe, "Vision-based human motion analysis: An overview," *Computer vision and image understanding*, vol. 108, pp. 4-18, 2007.
- [5] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *Acm computing surveys (CSUR)*, vol. 38, p. 13, 2006.
- [6] J. S. Lim and W. H. Kim, "Detection and tracking multiple pedestrians from a moving camera," in *Advances in Visual Computing*, ed: Springer, 2005, pp. 527-534.
- [7] J. Lim and W. Kim, "Detecting and tracking of multiple pedestrians using motion, color information and the AdaBoost algorithm," *Multimedia tools and applications*, vol. 65, pp. 161-179, 2013.
- [8] Y.-S. Hsieh, Y.-C. Su, and L.-G. Chen, "Robust moving object tracking and trajectory prediction for visual navigation in dynamic environments," in *Consumer Electronics (ICCE), 2012 IEEE International Conference on*, 2012, pp. 696-697.
- [9] B. Jung and G. S. Sukhatme, "Detecting moving objects using a single camera on a mobile robot in an outdoor environment," in *International Conference on Intelligent Autonomous Systems*, 2004, pp. 980-987.
- [10] A. Talukder, S. Goldberg, L. Matthies, and A. Ansar, "Real-time detection of moving objects in a dynamic scene from moving robotic vehicles," in *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, 2003, pp. 1308-1313.
- [11] Y.-K. Jung, K.-W. Lee, and Y.-S. Ho, "Feature-based object tracking with an active camera," in *Advances in Multimedia Information Processing—PCM 2002*, ed: Springer, 2002, pp. 1137-1144.
- [12] M. Siam, R. ElSayed, and M. ElHelw, "On-board multiple target detection and tracking on camera-equipped aerial vehicles," in *Robotics and Biomimetics (ROBIO), 2012 IEEE International Conference on*, 2012, pp. 2399-2405.
- [13] D. D. Doyle, A. L. Jennings, and J. T. Black, "Optical flow background estimation for real-time pan/tilt camera object tracking," *Measurement*, vol. 48, pp. 195-207, 2014.
- [14] A. Kundu, C. Jawahar, and K. M. Krishna, "Realtime moving object detection from a freely moving monocular camera," in *Robotics and Biomimetics (ROBIO), 2010 IEEE International Conference on*, 2010, pp. 1635-1640.
- [15] A. Vavilin, L.-M. Ha, and K.-H. Jo, "Camera motion estimation and moving object detection based on local feature tracking," in *Advanced Research in Applied Artificial Intelligence*, ed: Springer, 2012, pp. 544-552.
- [16] M. Zaki and M. Youssef, "TNRAC: a system for tracking multiple moving non-rigid objects using an active camera,"

- Signal, image and video processing, vol. 3, pp. 145-155, 2009.
- [17] B. Zitova and J. Flusser, "Image registration methods: a survey," Image and vision computing, vol. 21, pp. 977-1000, 2003.
- [18] D. I. Barnea and H. F. Silverman, "A Class of Algorithms for Fast Digital Registration," IEEE Transactions on Computers, vol. 21, pp. 179-186, 1972.
- [19] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching," DTIC Document1977.
- [20] A. Goshtasby, "Image registration by local approximation methods," Image and Vision Computing, vol. 6, pp. 255-261, 1988.
- [21] A. Rosenfeld and A. C. Kak, Digital picture processing vol. 1: Elsevier, 2014.
- [22] G. Stockman, S. Kopstein, and S. Bennett, "Matching images to models for registration and object detection via clustering," Pattern Analysis and Machine Intelligence, IEEE Transactions on, pp. 229-241, 1982.
- [23] F. Li, G. Zhang, and J. Yan, "Coregistration based on SIFT algorithm for synthetic aperture radar interferometry," in Proceedings of ISPRS Congress, 2008.
- [24] L. Xiaoming, Z. Lian, and H. Zhanyi, "SIFT based automatic registration of remotely-sensed imagery," Journal of remote sensing, vol. 10, pp. 885-892, 2006.
- [25] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform robust scale-invariant feature matching for optical remote sensing images," Geoscience and Remote Sensing, IEEE Transactions on, vol. 49, pp. 4516-4527, 2011.
- [26] A. Sedaghat and H. Ebadi, "Remote sensing image matching based on adaptive binning SIFT descriptor," Geoscience and Remote Sensing, IEEE Transactions on, vol. 53, pp. 5283-5293, 2015.
- [27] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," Communications of the ACM, vol. 24, pp. 381-395, 1981.
- [28] J. L. Mundy and A. Zisserman, Geometric invariance in computer vision vol. 92: MIT press Cambridge, MA, 1992.
- [29] N. Otsu, "A threshold selection method from gray-level histograms," Automatica, vol. 11, pp. 23-27, 1975.
- [30] A. Hakeem, K. Shafique, and M. Shah, "An object-based video coding framework for video sequences obtained from static cameras," in Proceedings of the 13th annual ACM international conference on Multimedia, 2005, pp. 608-617.
- [31] K. Gupta and A. V. Kulkarni, "Implementation of an automated single camera object tracking system using frame differencing and dynamic template matching," in Advances in Computer and Information Sciences and Engineering, ed: Springer, 2008, pp. 245-250.
- [32] R. Kasturi, D. Goldgof, P. Soundararajan, V. Manohar, M. Boonstra, and V. Korzhova, "Performance evaluation protocol for face, person and vehicle detection & tracking in video analysis and content extraction (VACE-II)," Computer Science & Engineering University of South Florida, Tampa, 2006.



## **Compensating surveillance camera movements using sequential image registration for car detection**

Ali Karami <sup>1\*</sup>, Mohsen Soryani <sup>2</sup>, Masood Varshosaz <sup>3</sup>

1- Ms.c student of remote sensing & photogrammetry, Geodesy & Geomatics Engineering Faculty, K.N.Toosi University of Technology

2- Associate professor in Department of Computer Vision, Computer Engineering Faculty, Iran University of Science & Technology

3- Associate professor in Department of remote sensing & photogrammetry, Geodesy & Geomatics Faculty, K.N.Toosi University of Technology

### **Abstract**

Recognizing and detecting cars in videos is ones of the main issues in computer visions. The main assumption of surveillance cameras is that the camera is fixed while taking videos. Any movements due to wind or external forces causes frequent vibrations in the camera and shift of image pixels. This makes conventional methods to detect static objects as moving ones mistakenly. The main aim of this research is to register sequential frames with background image to remove the noise and the mentioned problem. The algorithm first extracts sequential frames and background image using median method, and then registers all frames with the background image. Cars are detected by subtracting background image from frames. Three data were used to evaluate the capability of the proposed methods. Each data set has different car density and contain frequent vibrations while recording the videos. To evaluate the proposed method, FAR, MODP, HR and error percentage criteria were calculated with and without registration. The overall accuracies were 89% and 76% respectively which shows the 13% improvement in the accuracy of detecting cars with registration technique.

**Key words:** registration, object recognition and detection, matching, moving objects, feature extraction

---

**Correspondence Address** Tehran - Faculty of geodesy and geomatric - K.N.Toosi University of technology.  
Tel : 09187839086  
Email : ali.karami.surveying@gmail.com