نشریه علمی پژوهشی مهندسی فناوری اطلاعات مکانی

سال پنجم، شماره چهارم، زمستان ۱۳۹۶

Vol.5, No.4, Winter 2018

۲۳-٤۰

جبران حرکت دوربین‌های کنترل نظارتی با استفاده از ثبت تصاویر متوالی در شناسایی خودروها

علی کرمی [۱*]، محسن سریانی [۲]، مسعود ورشوساز [۳]

۱-دانشجوی کارشناسی ارشد فتوگرامتری، دانشکده نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی

۲-دانشیار گروه هوش مصنوعی ، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران

۳-دانشیار گروه فتوگرامتری و سنجش از دور، دانشکده نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی

تاریخ دریافت مقاله: ۱۳۹۵/۰۳/۱۸ تاریخ پذیرش مقاله: ۱۳۹۵/۱۰/۲۸

چکیده

تشخیص خودروها در تصاویر ویدئویی به‌عنوان یک موضوع کلیدی در مباحث بینایی کامپیوتر محسوب می‌گردد. فرض اصلی اکثر دوربین‌های نظارتی، ثابت بودن دوربین در حین فیلم‌برداری می‌باشد ولی اگر دوربین نظارتی به هر دلیلی مثلا براثر باد یا فشارهای بیرونی محیط اطراف دچار لرزش‌های مکرر شود دیگر روش‌های موجود برای شناسایی خودروها به‌درستی عمل نمی‌کنند. این موضوع باعث جابجا شدن پیکسل‌های تصویر و ایجاد نویز می‌شود که در اکثر مواقع باعث شناسایی عوارض ثابت به‌عنوان عوارض متحرک می‌گردد. هدف اصلی در این تحقیق استفاده از روش ثبت(هم‌مرجع سازی) فریم‌های متوالی با تصویر پس‌زمینه، برای حذف نویز و مشکل یادشده می‌باشد. روند کار به این صورت می‌باشد که بعد از استخراج فریم‌های متوالی و تصویر پس‌زمینه از طریق روش میانه‌گیری، تمام فریم‌های موجود با تصویر پس‌زمینه ثبت شده و سپس با استفاده از تفریق پس‌زمینه به شناسایی خودروها پرداخته می‌شود. سه نوع داده برای پیاده‌سازی و ارزیابی توانایی روش پیشنهادی مورد استفاده قرار گرفته است. ویژگی اصلی این مجموعه داده‌ها لرزش‌های مکرر دوربین در حین فیلم‌برداری و متفاوت بودن تراکم خودروهای موجود در هر مجموعه می‌باشد. از معیارهای HR (نرخ تشخیص درست)، MODP (دقت کشف چند شیئی)، FAR (نرخ تشخیص اشتباه) و OA (شاخص دقت کلی) برای ارزیابی روش پیشنهادی در دو حالت با ثبت تصویر(روش پیشنهادی) و بدون ثبت تصویر استفاده شده است. به طورکلی دقت به‌دست‌آمده برای شناسایی خودروها با ثبت تصاویر و بدون ثبت تصاویر به‌طور متوسط به‌ترتیب ۸۹ درصد و ۷۶ درصد می‌باشد. بنابراین می‌توان با استفاده از روش ثبت تصاویر حدود ۱۳درصد دقت شناسایی خودروها را افزایش داد.

کلید واژه‌ها: ثبت تصاویر، شناسایی عوارض، تناظریابی، عوارض متحرک.

* نویسنده مکاتبه کننده: تهران – دانشکده نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

تلفن: ۰۹۱۸۷۸۳۹۰۸۶

Email : ali.karami.surveying@gmail.com

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]

[DOI: 10.29252/jgit.5.4.23]

۱- مقدمه

تشخیص خودروها در تصاویر ویدئویی به‌عنوان 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 به‌منظور 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 به‌درستی عمل نمی‌کنند. دلیل این موضوع آنست که پیکسل‌های تصویر جابه‌جا می‌شوند که این خود در اکثر مواقع باعث شناسایی عوارض ثابت به‌عنوان عوارض متحرک می‌شود[۱ و ۵].

تاکنون روش‌های مختلفی برای جبران حرکت دوربین نظارتی انجام شده است. به‌عنوان مثال لیم[۱] سال‌های ۲۰۰۵ و ۲۰۱۳ در [۶ و ۷] و همچنین چن[۲] 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. به‌طور کلی مشکل اصلی این روش‌ها این است که دقت روش به ساختار داخلی دوربین مانند دیلای (Delay[۳]) 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) درآن صورت اختلافات ناشی از حرکات اضافی دوربین قابل تشخیص و حذف می‌باشند و همچنین خطای دیلای در شناسایی خودروها اثر گذار نخواهد بود. بنابراین در این مقاله از پس‌زمینه به‌عنوان 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 به‌طور 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 متدولوژی نحوه جبران حرکت تشریح 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 نهایتا در بخش هشتم پیشنهاداتی برای تحقیقات آینده خواهیم داشت.

۲- پیشینه تحقیق

تاکنون تحقیقات زیادی به‌منظور 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[۴] 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 آوردند و بردار جابجایی آن بلوک را به عنوان بردارحرکت دوربین درنظر گرفتند. در [۹] ردیابی اشیاء با دوربین تعبیه شده روی یک ربات متحرک در محیط‌های بیرونی بررسی شده است. در این کار برای جبران

[1] Lim
[2] Chen
[3] . تاخیر احتمالی در تصویر برداری

[4] Block Matching Algorithm

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

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[1] استفاده شده است. نویسندگان در این مقاله از جریان نوری برای کشف حرکت پیکسل‌ها در بین فریم‌های متوالی و به‌دست آوردن بردار حرکت دوربین از طریق بردار جابجایی 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[2] ارائه شده است. در این کار نویسندگان با استفاده از تناظر ویژگی‌های گوشه به‌دست آمده با الگوریتم هریس[3] بین دو فریم متوالی حرکت دوربین را به‌دست می‌آورند سپس با اعمال حرکت دوربین به فریم t-1 و تفاضل این فریم از فریم t، ناحیه مربوط به عابرین پیاده مشخص می‌گردد. در [۱۱] ردیابی اشیاء متحرک بر اساس ویژگی با استفاده از یک دوربین دارای حرکات انتقالی[4]، مورد بررسی قرار می‌گیرد. در این روش نویسندگان 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 صرفنظر کنند. ایراد اصلی این روش محدود کردن حرکت دوربین به حرکت انتقالی می‌باشد. در [۱۲] نیز از ویژگی‌های گوشه به‌دست آمده با الگوریتم فست[5] برای جبران حرکت دوربین [6] استفاده شد. در این روش موقعیت اشیاء متحرک با استفاده از یک الگوریتم کشف حرکات پرت به‌دست 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 به‌صورت یک شبکه از مثلث‌ها ساختاربندی شده و از رئوس این مثلث‌ها به‌عنوان ویژگی بین فریم‌های متوالی استفاده می‌شود. سپس در فریم بعدی ویژگی‌های متناظر را پیدا کرده و بر اساس همسایگی این نقاط متناظر شبکه را به‌روز رسانی می‌کنند. سرانجام با استفاده از این نقاط متناظر پارامترهای حرکت دوربین را به‌دست آورده و بردارهایی که به پس‌زمینه تعلق ندارند را به‌عنوان 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 پرت (نقاط ویژگی مرتبط با اشیاء متحرک) حرکت پس‌زمینه 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 t و t+1 عمل شده است. اگر دوربین مورد استفاده خطای تاخیر احتمالی داشته باشد، دقت روش‌های یاد شده برای جبران حرکت تحت تاثیر این خطا قرار می‌گیرد. بنابراین هدف اصلی در این تحقیق ارائه روشی جهت این مسئله می‌باشد. برای این کار اگر یک سیستم مختصات مرجع را ثابت درنظر 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 کاملا مستقل از این خطا عمل می‌کند. تنها خطایی که باقی می‌ماند تغییر در موقعیت مکانی

[1] Optical Flow
[2] AdaBoost
[3] Hariss
[4] Pan Tilt Zoom
[5] FAST
[6] Compensating Camera Movements

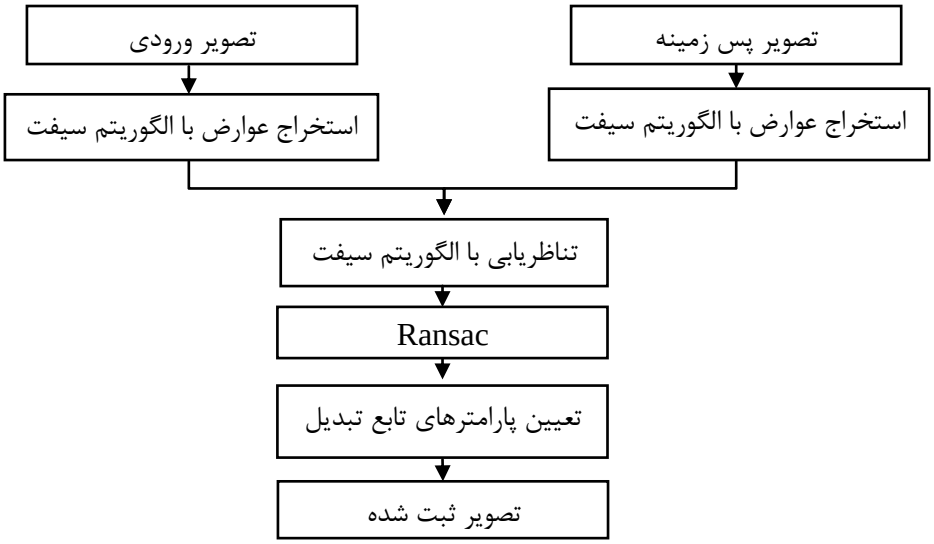
[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

پیکسل‌ها می‌باشد که با استفاده از ثبت تصاویر ورودی به تصویر پس‌زمینه(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 پس‌زمینه

ثبت تصویر روند تعیین بهترین تطبیق مکانی بین دو یا چند تصویر به‌دست‌آمده در زمان‌های مختلف، از منظرهای متفاوت و یا با سنسورهای مختلف می‌باشد که دو تصویر مبنا و ورودی را به‌صورت هندسی بر هم منطبق می‌نماید[۱۷]. در سال‌های اخیر تحقیقات گسترده‌ای درزمینه تعیین اتوماتیک تابع تبدیل ریاضی برای ثبت تصاویر انجام پذیرفته است[۱۸، ۱۹، ۲۰، ۲۱، ۲۲].

پیچیدگی جستجوی تشابهات، زمان و حجم بالای محاسبات داده‌های تصویری از مشکلات اصلی در این زمینه هستند. لذا احتیاج به روش‌هایی که بتوانند راه‌حل 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 موردنظر، باید روشی را طراحی نمود تا بتواند نه‌تنها خطاها و 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 داده‌شده‌اند، طراحی شود. در ادامه به تشریح مراحل الگوریتم پیشنهادی پرداخته می‌شود.



شکل ۱: فلوچارت پیشنهادی برای ثبت تصاویر

۳-۱- استخراج عوارض

اولین مرحله برای ثبت دو تصویر، استخراج عوارض(ویژگی‌ها) می‌باشد. عوارضی که تاکنون برای برقراری ارتباط دو تصویر استفاده گردیده‌اند بسیار متنوع می‌باشند و روش‌های تناظریابی بر این اساس به دو دسته اصلی تقسیم می‌شوند که یا عارضه مبنا و یا ناحیه مبنا هستند[۲۳]. الگوریتم سیفت یکی از روش‌های موجود برای استخراج عوارض متناظر بوده و غالبا برای کاربردهایی نظیر تشخیص شی، تناظریابی تصاویر، ردیابی، بازیابی شی در بانک‌های اطلاعاتی

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

چندرسانه‌ای، ربات‌های خودکار محلی و غیره مورداستفاده 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 به‌طور موفقیت‌آمیزی تعداد زیادی عوارض نقطه‌ای پایدار را استخراج کرد[۲۴]. در مقاله حاضر نیز از الگوریتم سیفت برای استخراج عوارض استفاده‌شده است.

۳-۲- تناظریابی

در این مرحله مقایسه یک بردار توصیفگر از تصویر مرجع با تمام بردارهای توصیفگر نقاط کلیدی در تصویر هدف، به‌صورت ضرب داخلی دو بردار انجام می‌پذیرد. مفهوم ضرب داخلی در اینجا بر اساس آنالیز برداری، همان مفهوم حداقل فاصله اقلیدسی را دارد. از نتیجه یک‌ضرب داخلی بردارهای توصیفگر دو تصویر، یک ماتریس برداری n×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 به‌عنوان دو نقطه متناظر انتخاب می‌کنیم [۲۵ و ۲۶].

۳-۳- حذف نقاط اشتباه

به علت اینکه بعضی از انطباق‌های اشتباه در مرحله تناظریابی تصاویر ممکن است رخ دهد، با استفاده از الگوریتم رنسک [1] و در قالب مدل هموگرافی، صحت انطباق جفت تناظرهای اولیه بررسی‌شده و انطباق‌های غیر صحیح حذف می‌شوند[۲۷]. این الگوریتم توسط فیشلر و بولس [2] در سال ۱۹۸۱ معرفی شد[۲۷]. الگوریتم رنسک روشی نیرومند به‌منظور برازش یک مدل هندسی به مجموعه‌ای از موقعیت‌های متناظر است که تعدادی جفت‌های متناظر اشتباه نیز در میان آن‌ها وجود دارد[۲۷].

رنسک فرض می‌کند که با مجموعه‌ای از داده‌های صحیح [3] می‌توان پارامترهای مدل را به‌گونه‌ای تخمین زد که به‌طور بهینه به این داده‌ها برازش شود. داده‌های پرت [4] از مقادیر زیاد نویز یا از اندازه‌گیری‌های غلط یا فرض‌های نادرست در تفسیر داده رخ می‌دهند و نمی‌توانند به مدل، برازش شوند. این تخمین با احتمال مشخصی صحیح است، زیرا رنسک یک تخمین زننده اتفاقی می‌باشد. این الگوریتم به‌طور وسیع در مسائل تخمین پارامترهای مدل در بینایی کامپیوتری از قبیل تناظریابی عوارض، و آشکارسازی هندسی کاربرد دارد.

۳-۴- 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، $T: R^2 \to R^2$ می‌تواند به‌صورت ماتریس 2*2 غیرسینگولار A، و بردار انتقال $t \in R^2$، بیان گردد.

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

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

[1] Ransac
[2] Fischler , Bolles
[3] Inlier
[4] Outlier

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

مقیاس و دوران و عمود نبودن محورها می‌باشد[۲۸].

۴- تشخیص خودروها

برای تشخیص خودروها به این صورت عمل می‌شود که بعد از اینکه فریم‌ها استخراج شدند، از بین این تصاویر، تصویر پس‌زمینه [۱] محاسبه خواهد شد و کل تصاویر ورودی از پس‌زمینه کم می‌کنیم تا عوارض متحرک از تصویر استخراج شوند. زمانی که دوربین دارای حرکات اضافی باشد الگوریتم‌هایی که از مدل سازی پس‌زمینه یا تفاضل فریم برای به‌دست آوردن ناحیه‌های مرتبط به اشیاء متحرک استفاده می‌کنند ناحیه‌های نادرستی(عوارض ثابت) را از پس‌زمینه به‌عنوان اشیاء متحرک برچسب گذاری می‌کنند. همچنین ناحیه‌های نادرستی را از پس‌زمینه به‌عنوان اشیاء متحرک برچسب گذاری می‌کنند که در واقع ثابت هستند و حرکت دوربین باعث ایجاد خطا در این موارد می‌شود. برای جبران این حرکات اضافی دوربین بعد از استخراج پس‌زمینه تمام فریم‌های جاری را نسبت به پس‌زمینه ثبت می‌کنیم و بعد با استفاده از روش تفاضل فریم‌ها به شناسایی خودروها پرداخته می‌شود. در شکل (۲) فلوچارت پیشنهادی برای شناسایی خودروها نشان داده‌شده است.

برای استخراج خودروها، همان‌طور که در شکل (۲) نشان داده‌شده است، بعدازاینکه تصاویر از فیلم ویدیویی استخراج شد اگر یک پس‌زمینه ثابت نداشته باشیم که عاری از خودروها و عوارض متحرک دیگر باشد باید یک تصویر پس‌زمینه از فریم‌های ویدیو استخراج کرد.

پس‌زمینه ممکن است به میزان زیادی توسط واقع‌شدن سایه ساختمان‌ها یا ابرها و یا به‌دلیل تغییر در شرایط روشنایی صحنه، تغییر یابد. با این شرایط محیطی متغیر، تصویر پس‌زمینه باید به‌طور منظم به‌هنگام شود. چندین روش برای به‌هنگام‌سازی تصویر پس‌زمینه وجود دارد. متداول‌ترین آن‌ها میانگین‌گیری و به هنگام‌سازی انتخابی است. در میانگین‌گیری، پس‌زمینه به شیوه‌ای تدریجی [۲] با میانگین‌گیری از پس‌زمینه قبلی و فریم جاری، ساخته می‌شود. اگر یک میانگین وزن‌دار را بر پس‌زمینه قبلی و فریم جاری اعمال کنیم، تصویر پس‌زمینه به روش به‌هنگام‌سازی انتخابی ساخته‌شده است.

روش دیگر استخراج تصویر پس‌زمینه استفاده از فیلتر میانه [۳] است. لازم به ذکر است که هریک از روش‌های ذکرشده دارای مزایا و معایبی می‌باشند. به‌عنوان‌مثال ازآنجاکه میانگین شدت رنگ مجموعه‌ای از پیکسل‌های تشکیل‌دهنده 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 تشکیل‌دهنده 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 پس‌زمینه استفاده‌شده است.

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

بعدازاین که عملیات ثبت روی فریم‌ها صورت گرفت برای شناسایی عوارض متحرک، هر تصویر ورودی با تصویر پس‌زمینه تفاضل مطابق رابطه (۲) گرفته می‌شود.

$$Dif_n = |updated\ background_p - F_n| \quad (۲)$$

$updated\ background_p$ تصویر پس‌زمینه به‌روزرسانی مرحله p ام، و F_n تصویر هم‌مرجع شده n ام می‌باشد.

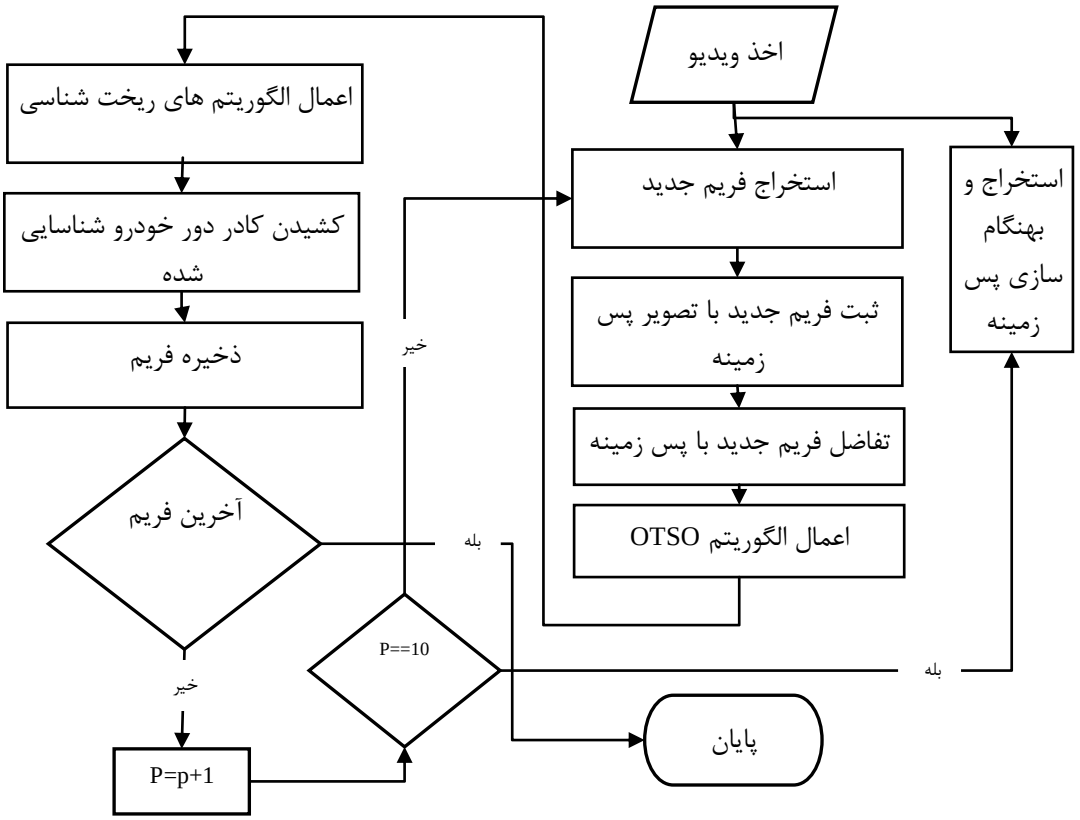
[1] Background

[2] Gradual

[3] Median Filter

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]

[DOI: 10.29252/jgit.5.4.23]



شکل ۲: الگوریتم پیشنهادی برای شناسایی خودروها

در تصویر تفاضل ممکن است نواقصی در خودروهای شناسایی شده وجود داشته باشد، به‌عنوان‌مثال 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 به‌خصوص در مرحله تشخیص خودروها شوند. بنابراین در ادامه کار از روش‌های ریخت‌شناسی[1] موجود در حوزه پردازش تصویر جهت برطرف کردن این نواقص بهره خواهیم گرفت. مراحل به‌کار گرفته‌شده 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 دوسطحی اعمال خواهد شد.

روش ارائه‌شده توسط اتسو[2] که یکی از روش‌های معروف در آستانه‌گیری سراسری است[3] و بر اساس هیستوگرام تصویر عمل می‌نماید، کاربرد فراوانی در پردازش تصویر دارد. در این پژوهش از روش اتسو استفاده‌شده 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.

[1] Morphology

[2] Otsu
[3] Global thresholding

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

همچنین این عملگر منجر به پر شدن فضای خالی ایجادشده در میان اشیای استخراج‌شده می‌شود.
پس از اعمال عملگر بستن، مؤلفه‌های همبند را از درون تصویر استخراج می‌کنیم. درواقع مؤلفه‌های همبند، نمایشی از ناحیه اشغال‌شده توسط خودروها در تصویر می‌باشند که این ناحیه یک همسایگی را در فضای تصویر تعریف می‌کند. سپس برای به‌دست آوردن مراکز تقریبی خودروها، بر روی هریک از خودروهای استخراج‌شده مستطیلی محاطی رسم می‌کنیم و سپس مرکز این مستطیل به‌عنوان مرکز ثقل یا موقعیت خودرو استخراج می‌شود. با استفاده از مستطیل‌های محاطی، دور خودروهای شناسایی‌شده در فریم اصلی یک پنجره کشیده می‌شود.

۵- پیاده‌سازی

برای پیاده‌سازی الگوریتم پیشنهادی از نرم‌افزار متلب استفاده‌شده 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.

جدول ۱: داده‌های مورد استفاده برای ارزیابی سیستم پیشنهادی

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

ویژگی اصلی این پایگاههای داده آن است که دوربین نظارتی براثر باد و فشارهای جانبی دچار لرزش‌های متوالی شده که در تشخیص خودروها می‌تواند مشکل ایجاد کند.
ویدئوهای موردنظر را با استفاده از امکانات ارائه‌شده در جعبه‌ابزار پردازش ویدئویی نرم‌افزار متلب، به فریم‌های متوالی تبدیل کردیم تا پردازش‌های بعدی را بر روی این تصاویر استخراج‌شده انجام دهیم. شکل (۳) نمونه‌ای از فریم‌های استخراج‌شده را نشان می‌دهد.

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

شکل ۳: نمونه‌ای از فریم‌های استخراج‌شده

برای استخراج تصویر پس‌زمینه در این تحقیق از فیلتر میانه به دلیل حفظ جزئیات موجود در تصویر استفاده شده است. شکل (۴) نمونه‌ای از تصاویر پس زمینه استخراج شده توسط این روش برای سه ویدیویی مورد بررسی را نشان می‌دهد.

شکل ۴: نمونه‌ای از تصاویر پس‌زمینه استخراج شده برای سه ویدیو

۵-۱- جبران حرکت از طریق ثبت تصاویر به پس‌زمینه

پس از استخراج تصویر پس‌زمینه تمام فریم‌های ورودی را یک به یک با تصویر پس‌زمینه ثبت می‌کنیم. برای تصویر پس‌زمینه و هر فریم ورودی با الگوریتم سیفت تناظریابی صورت گرفته و بردارهای ویژگی به‌دست می‌آیند. از طریق این بردار ویژگی و همچنین یک تابع تبدیل افاین می‌توان این تصاویر را بر روی هم ثبت کرد. در شکل (۵) نتیجه ثبت تصاویر با پس‌زمینه نشان داده‌شده است.

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

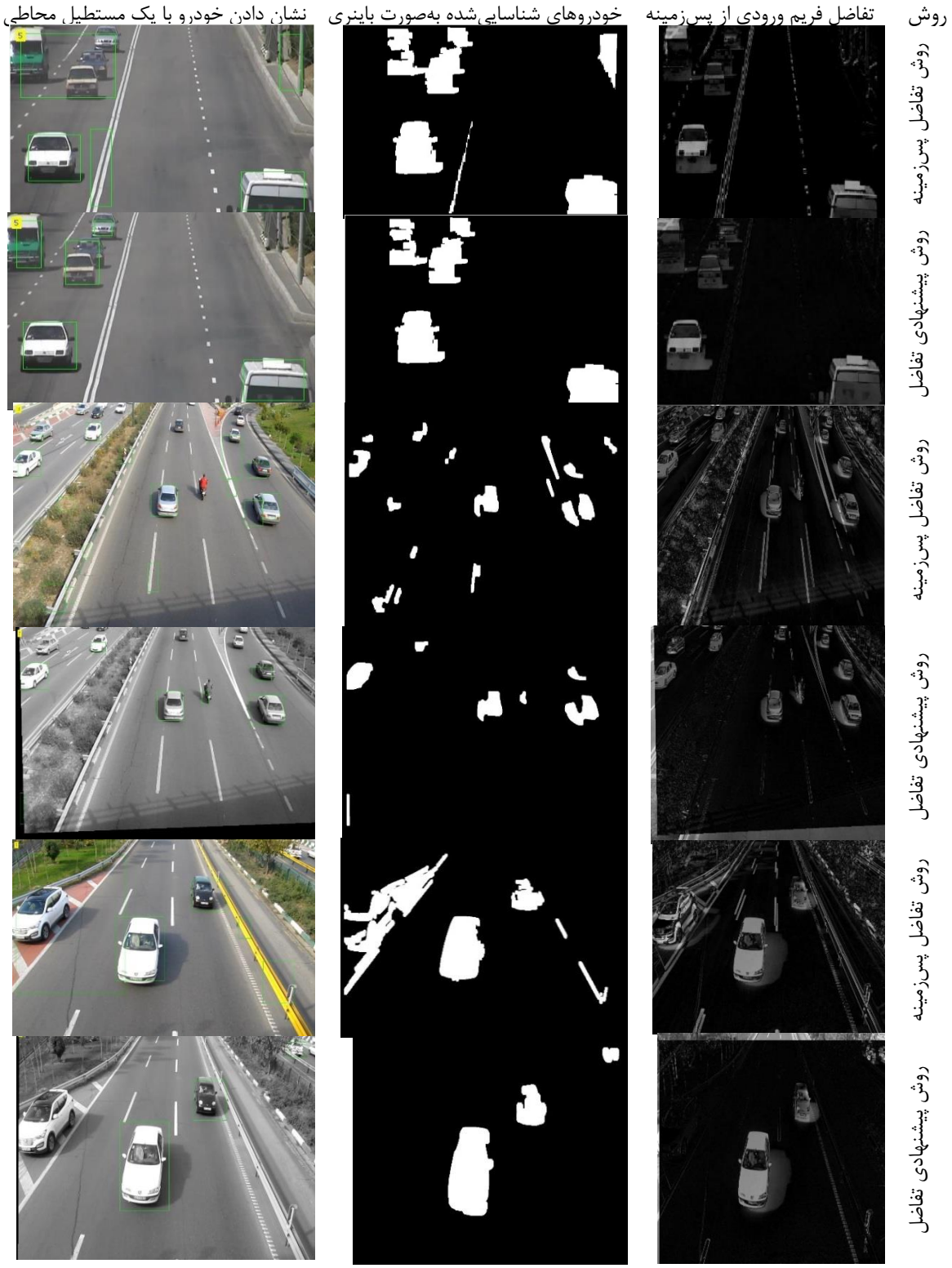
شکل ۵ : سطراول) فریم ورودی. سطردوم) میزان تغییرات فریم ورودی نسبت به پس‌زمینه. سطرسوم) فریم ثبت شده با تصویر پس‌زمینه

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 پس‌زمینه را نشان می‌دهد. در این اشکال همانطور که نشان داده شده است، پیکسل‌های که نسبت به پس‌زمینه دارای تغییر می‌باشد به‌صورت رنگی نشان داده شده است. سطر سوم نتایج ثبت تصاویر ورودی با پس‌زمینه را نشان می‌دهد.

۵-۲- استخراج خودروها

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

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]



شکل ۶ : ستون اول) تفاضل فریم‌های متوالی از تصویر پس‌زمینه در هر دو حالت با ثبت و بدون ثبت. ستون دوم) خودروهای شناسایی‌شده به‌صورت 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

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

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. درستون دوم خودروهای شناسایی‌شده به‌صورت باینری در هر دو حالت با ثبت و بدون ثبت دیده می شوند. 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 دوسطحی، اعمال خواهد شد. برای اعمال این آستانه از روش رایج اتسو که آستانه بهینه را پیدا می‌کند استفاده‌شده است. برای حذف نواحی کوچک حاصل از نویزها از الگوریتم‌های ریخت شناسی[1] مانند بازکردن و بستن با ماسک ۹*۹ استفاده‌شده است(شکل(۶) ستون دوم). پس برای به‌دست آوردن دقیق‌تر مراکز خودروها، بر روی هریک از خودروهای استخراج‌شده به‌صورت اتوماتیک مستطیلی محاطی رسم شده و سپس مرکز این مستطیل محاطی به‌عنوان مرکز ثقل یا موقعیت خودرو استخراج می‌شود(شکل(۶) ستون سوم).

همانطور که در سطرهای فرد شکل(۶) پیداست به‌دلیل حرکات اضافی که دوربین در حین فیلم‌برداری داشته است، دو خطا 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 به‌عنوان یک خودرو شناسایی شوند. دوم اینکه ناحیه‌های نادرستی(عوارض ثابت) از پس‌زمینه بعنوان اشیاء متحرک برچسب‌گذاری می‌شوند. ولی با اعمال ثبت تصاویر حرکات اضافی 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 خطا می‌باشد. یک ماتریس ابهام، یک آرایه مربع با ابعاد $n\times n$ می‌باشد، که n تعداد کلاس‌ها را نشان می‌دهد و دراین 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 می‌دهدکه به‌صورت دستی خودروها استخراج شده‌اند. مجموعه دوم، از برچسب پیکسل‌هایی تشکیل شده است که به‌وسیله روش پیشنهادی استخراج شده‌اند جدول (۲).

[1] Morphology

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]

[DOI: 10.29252/jgit.5.4.23]

جدول ۲: ماتریس ابهام

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

ستون‌های ماتریس ابهام داده مرجع را نشان می‌دهد، و سطرهای آن، برچسب‌های اختصاص داده شده توسط روش پیشنهادی را نشان میدهد. در این ماتریس TP[1] تعداد نقاطی است که در هر دو مجموعه داده‌ها برچسب توصیفی برای آن شناسایی شده است، در حالی که TN[2] تعداد نقاطی را مشخص می‌کند که در هر دو داده مورد شناسایی واقع نشده باشد. همچنین FP[3] تعداد نقاطی هستند که توسط روش پیشنهادی به اشتباه شناسایی شده‌اند. در نهایت FN[4] تعداد نقاطی است که توسط روش‌پیشنهادی به اشتباه مورد شناسایی واقع نشده است.

۱) شاخص دقت کلی[5]

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 به‌درستی به کلاس خودرو تعلق گرفته‌اند به تعداد کل پیکسل‌ها یا خودروها محاسبه می‌شود[۳۰, ۳۱]

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[6]

برای هر کلاس اطلاعاتی در ماتریس ابهام، نرخ تشخیص درست به وسیله تقسیم درایه (i,j) به مجموع ستون i، محاسبه می‌شود. نرخ تشخیص درست رابطه (۴)، در حقیقت، معیاری از شاخص جامع بودن[7](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.

رابطه (۴)

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

۳) نرخ تشخیص اشتباه[8]

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].

رابطه (۵)

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

۴) معیار دقت کشف چند شیئی MODP[9]

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.

رابطه (۶)

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

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

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

[1] True Positive
[2] True Negative
[3] False Positive
[4] False Negative
[5] Overall Accuracy
[6] Hit Rate
[7] Completeness
[8] False Alarm Rate
[9] Multi Object Detection Precision

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

استخراج شده iام در فریم tام می‌باشند. همچنین N^t_{mapped} مشخص کننده تعداد عارضه تصویر شده در فریم tام 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 پس‌زمینه استخراج می‌شود. سپس با استفاده از روش تفاضل پس‌زمینه تمام فریم‌های متوالی را از پس‌زمینه کم کرده و عوارض متحرک استخراج‌شده است. حالت دوم بعداز اینکه تصویر پس‌زمینه استخراج شد، تمام فریم‌های متوالی با تصویر پس‌زمینه ثبت شده و بعد تفاضل گرفته می‌شود.

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

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

نام داده	تعداد فریم	*تعدادکل* خودروها	تشخیص صحیح	تشخیص غلط(تلفیق و برعکس)	غیرقابل تشخیص	تشخیص عوارض ثابت به‌عنوان خودرو	دقت کلی (درصد)	روش کار
۱MVI_000	۵۰۰	۴۴۳۰	۳۹۵۰	۴۶۵	۱۵	۵۲	۸۸٫۱۳	با ثبت
			۳۳۴۵	۷۵۷	۳۱	۴۲۵	۶۸٫۸۹	بدون ثبت
۲MVI_000	۵۰۰	۸۵۲	۷۸۹	۶۲	۱	۶	۹۱٫۹۵	با ثبت
			۷۶۲	۸۸	۲	۸۰	۸۱٫۷۵	بدون ثبت
Highway1	۵۰۰	۱۷۲۷	۱۵۴۶	۱۷۶	۵	۲۸	۸۸٫۰۹	با ثبت
			۱۴۶۶	۲۵۶	۵	۱۴۲	۷۸٫۴۴	بدون ثبت
جمع کل	۱۵۰۰	۷۰۰۹	۶۲۸۵	۷۰۳	۲۱	۸۶	۸۹٫۳۹	با ثبت
			۵۵۷۳	۱۱۰۱	۳۸	۶۴۷	۷۶٫۳۶	بدون ثبت
قدرمطلق تفاضل حالت با ثبت و بدون ثبت			۷۱۲	۲۹۸	۱۷	۵۶۱	۱۳٫۰۳	

مطابق جدول (۳) تعداد کل خودروهای موجود در تمام فریم‌ها، تعداد خودروهای شناسایی شده صحیح در دو حالت، تعداد خودروهایی که به اشتباه ادغام و یا تجزیه شده‌اند، تعداد خودروهای غیرقابل تشخیص و در نهایت تعداد عوارض ثابتی که به اشتباه به عنوان عارضه متحرک شناسایی شده، نشان داده شده است. براساس این اطلاعات برای هر کدام از داده‌ها مطابق رابطه (۳) دقت 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 اولاً تعداد خودروهای صحیح تشخیص داده‌شده قبل از اعمال الگوریتم پیشنهادی ۵۵۷۳ و بعد از اعمال الگوریتم پیشنهادی ۶۲۸۵ بوده است. که باعث تشخیص ۷۱۲ خودرو بیشتر از حالت قبل است.

در روشی که از ثبت تصویر استفاده نمی شود هنگامی که چندین خودرو کنار هم یا نزدیک به هم قرار گرفته‌اند در اثر لرزش‌هایی که به دوربین واردشده باعث شده که خودروهای نزدیک به هم باهم تلفیق شوند و چند خودرو را به‌عنوان یک خودرو شناسایی کند و هر چقدر تراکم خودروها زیاد باشد با

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

لـــرزش دوربـــین عمـــل تلفیـــق خودروهـــا بیشـــتر اتفـــاق مـــی‌افتـــد. عـــلاوه براینکـــه خودروهـــا را بـــاهم تلفیق می‌کند باعث می‌شود عـوارض ثـابتی کـه درجـه خاکستری آن‌ها با محیط زیاد فرق دارند(مثل خط‌کشی خیابان‌های شهری و ...) در دو فـریم متـوالی بـه‌عنـوان عـوارض متحـرک شناسـایی شـوند. درمجمـوع تعـداد عوارض ثابتی که به‌عنـوان عـوارض متحـرک شناسـایی شده، در حالت اول ۶۴۷ عارضـه مـی‌باشـد کـه بعـد از ثبت تصاویر این تعداد به ۸۶ عارضه کاهش‌یافتـه اسـت که این خود نشان‌دهنده توانایی ثبت تصاویر متـوالی بـا تصـویر پـس‌زمینـه در مـواقعی کـه دوربـین نظـارتی دچار لرزش‌های مکرر شده را نشان می‌دهد.

در ادامه بـرای بـه‌دسـت آوردن میـزان دقـت و قابلیـت اعتماد در تشخیص خودروها از معیارهای دیگری که در روابـط (۳، ۴، ۵ و ۶) بیـان شـده اسـتفاده مـی‌شـود. برای این کار به‌صورت اتفاقی ۱۰ فریم از هـر مجموعـه داده را انتخاب می‌کنیم. در دو حالـت بـدون ثبـت و بـا ثبت تصاویر، خودروها استخراج می‌شوند. پیکسـل‌هـای مربوط به موقعیـت دقیـق خودروهـا در ایـن ۱۰ فـریم به‌صورت دستی اسـتخراج شـده و بـه‌عنـوان داده‌هـای واقعی در ارزیابی مورد استفاده قـرار مـی‌گیرنـد. بعـد از این مرحله به تهیه ماتریس ابهام برای هـر ۱۰ فـریم در هر سه داده و در دو حالت مختلف با ثبت و بدون ثبـت پرداخته می‌شود.

به‌طور کلی بـرای ارزیـابی دقـت در ایـن تحقیـق از ۶۰ ماتریس ابهام(۱۰×۳×۲) استفاده شده اسـت. نتـایج در قالب نمودارهای شکل (۷) نشان داه شده است.

به‌طور کلی با توجه به نمودارهای شکل (۷) میـزان HR و MODP بعـد از ثبـت بیشـتر از حـالتی اسـت کـه عمل ثبت صورت نگرفته است و همچنین میـزان FAR هم بعد از ثبت نسبت به قبل از ثبت کمتر شده است و این خود نشان دهنده توانـایی الگـوریتم پیشـنهادی در شناسایی دقیق‌تر خودروها می‌باشد.

همان‌طـور کـه در جـدول (۳) نشـان داده شـده اسـت به‌دلیل تـراکم بـالای خودروهـای موجـود در داده‌هـای MVI_01، هم میزان خودروهایی کـه تلفیـق شـده‌انـد بیشتر بوده و هم اینکه تعداد عوارض ثابتی که به اشتباه بـه‌عنـوان عارضـه متحـرک تشـخیص داده شـده‌انـد، زیاد می‌باشد. در نتیجه تفاوت نمودارهای HR، MODP و FAR قبل از ثبت و بعد از ثبت برجسته می‌باشد. این نتایج نشان می‌دهد که هر چه حرکات اضـافی دوربـین حـین فـیلم‌بـرداری بیشـتر باشـد و همچنـین تـراکم خودروهـا بیشـتر باشـد، دقـت و قابلیـت اعتمـاد در شناسایی خودروها با استفاده از روش تفاضـل فـریم‌هـا کمتر می‌شود ولی اگر قبل از عمل تفاضل تمام فریم‌هـا را به پس‌زمینه ثبت کرد و بعد بـه اسـتخراج خودروهـا پرداخـت، درآن صـورت تـا حـد خـوبی دقـت و قابلیت اعتماد در شناسایی خودروها بالا می رود.

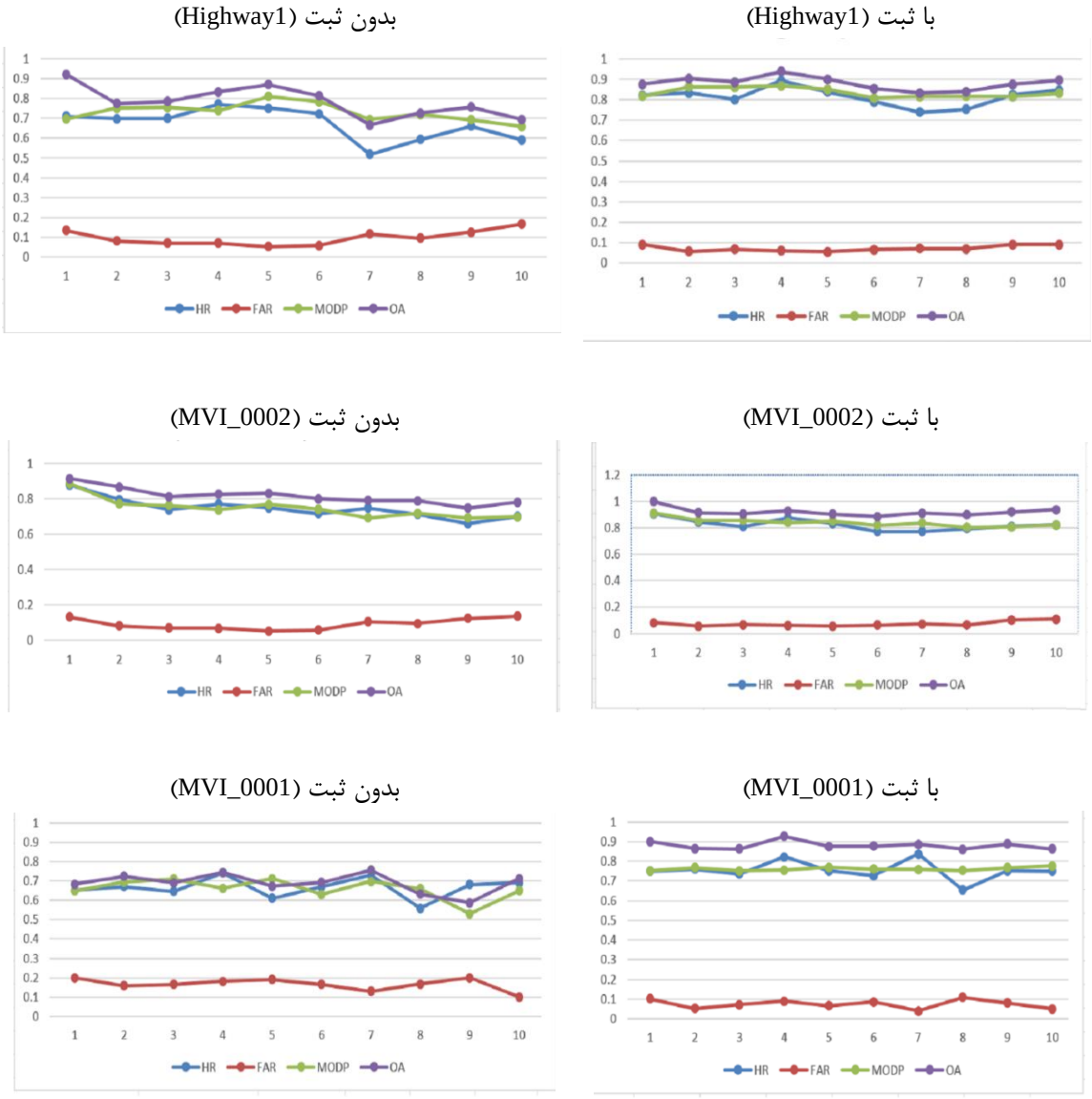
درمجمـــوع و بـــه‌طـــورکلی دقـــت نهـــایی(OA) تشخیص خودروها در حالت اول یعنی قبل از ثبت ۷۶٪ و با روش پیشنهادی یعنی بعد از ثبـت ۸۹٪ مـی‌باشـد. این نتایج نشان می‌دهد که عمل ثبت تصاویر متوالی بـا پـس‌زمینـه حـدود ۱۳٪ در تشـخیص خودروهـا بهتـر عمل کرده است.

۷-نتیجه‌گیری

در این تحقیـق روشـی مناسـب بـرای جبـران حرکـت اضافی دوربین‌های نظارتی در اثر عوامـل بیرونـی نظیـر باد و غیره با استفاده از ثبت تصاویر متوالی با پس‌زمینه در شناسایی خودروها، ارائه‌شده است. برای ایـن‌منظـور بعد از استخراج فریم‌ها و پیدا کردن تصویر پـس‌زمینـه با استفاده از روش میانـه‌گیـری، بایـد تمـام فـریم‌هـای ورودی با تصویر پس‌زمینه بر اساس الگوریتم پیشنهادی که در بخش(۳) ارائه‌شده، ثبت شوند. سپس با اسـتفاده از روش تفریق پس‌زمینه به شناسایی خودروها پرداخته شده است. این امر باعث حذف خطای دیـلای در حـین تصـویر بـرداری مـی‌شـود. بـرای ارزیـابی نتـایج، روش پیشـــنهادی بـــر روی ســـه نـــوع داده پیاده‌سازی شده‌است. ویژگی اصلی این مجموعه داده‌هـا لـرزش‌هـای مکـرر دوربـین در حـین فـیلم‌بـرداری و

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]
[DOI: 10.29252/jgit.5.4.23]

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 پرداخته‌شده است. در هر دو حالت با استفاده از متریک‌های MODP,HR,FAR و OA، مقایسه صورت گرفته است. در حالت اول تعداد عوارض ثابتی که به‌عنوان خودرو شناخته‌شده است حدود ۶۴۷ عارضه می‌باشد که با اعمال الگوریتم پیشنهادی 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 پیشنهای در شناسایی خودروها می‌باشد. نتایج به‌دست‌آمده در این تحقیق نشان داد که دقت کلی روش پیشنهادی نسبت به حالت اول(بدون عمل ثبت) حدود ۱۳درصد در شناسایی عوارض متحرک بیشتر شده است.



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

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]

[DOI: 10.29252/jgit.5.4.23]

۸- پیشنهادات

به‌منظور ادامه و تکمیل تحقیقات صورت گرفته در زمینه جبران حرکت اضافه دوربین پیشنهاد می‌گردد که با مقایسه روش‌های ثبت تصاویر به معرفی بهترین و کارآمدترین روش برای جبران حرکت اضافی دوربین با استفاده از روش تفاضل فریم با پس‌زمینه پرداخته شود.

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,"

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]

[DOI: 10.29252/jgit.5.4.23]

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. Barena 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. Benett, "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.

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]

[DOI: 10.29252/jgit.5.4.23]

K. N. TOOSI UNIVERSITY OF TECHNOLOGY
FACULTY OF GEODESY AND GEOMATICS ENGINEERING

Journal of Geospatial Information Technology
Vol.5, No.4, Winter 2018

Compensating surveillance camera movements using sequential image registration for car detection

Ali Karami [1]*, Mohsen Soryani [2], Masood Varshosaz [3]

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 geomatic - K.N.Toosi University of technology.
Tel : 09187839086
Email : ali.karami.surveying@gmail.com

[Downloaded from jgit.kntu.ac.ir on 2026-11-19]

[DOI: 10.29252/jgit.5.4.23]