



## مقایسه ویژگی‌های مختلف برای ثبت تصاویر به منظور جبران حرکت دوربین در روش تفاضل پس زمینه

علی کرمی<sup>۱\*</sup>، محمد شکری<sup>۲</sup>، محسن سریانی<sup>۳</sup>، مسعود ورشوساز<sup>۴</sup>

- ۱- دانشجوی کارشناسی ارشد فتوگرامتری، دانشکده نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی
- ۲- دانشجوی کارشناسی ارشد سنجش از دور، دانشکده نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی
- ۳- دانشیار گروه هوش مصنوعی، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران
- ۴- دانشیار گروه فتوگرامتری و سنجش از دور، دانشکده نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی

### چکیده:

تشخیص خودروها در تصاویر ویدئویی به عنوان یک موضوع کلیدی در مباحث تحلیل ترافیک محسوب می‌گردد. فرض اصلی اکثر دوربین‌های ناظریتی، ثابت بودن دوربین در حین فیلمبرداری می‌باشد ولی اگر دوربین ناظریتی به هر دلیلی مثلاً برادر باد یا فشارهای بیرونی محیط اطراف دچار لرزش‌های مکرر شود دیگر روش‌های موجود برای شناسایی خودروها به درستی عمل نمی‌کنند. این موضوع باعث جایجا شدن پیکسل‌های تصویر و ایجاد نویز می‌شود که در اکثر مواقع باعث شناسایی عوارض ثابت به عنوان عوارض متحرک می‌گردد. هدف اصلی در این تحقیق مقایسه ویژگی‌های مختلف برای ثبت (هم‌مرجع‌سازی) تصویر و ارائه بهترین و کارآمد ترین الگوریتم برای ثبت فریم‌های متوالی با تصویر پس زمینه جبران حرکت می‌باشد. روند کار به این صورت می‌باشد که برای هر کدام از ویژگی‌های مختلف ثبت، تمام فریم‌های موجود با تصویر پس زمینه ثبت شده و سپس با استفاده از تفربیق پس زمینه به شناسایی خودروها پرداخته می‌شود. از داده مربوط به یکی از اتوبان‌های کشور اکراین برای پیاده‌سازی و ارزیابی توانایی روش پیشنهادی استفاده شده است. ویژگی اصلی این مجموعه داده لرزش‌های مکرر دوربین در حین فیلمبرداری و متفاوت بودن تراکم خودروهای موجود در هر مجموعه می‌باشد. از معیارهای FAR, MODP و HR، برای ارزیابی روش پیشنهادی استفاده شده است. به طور کلی نتایج بدست آمده نشان بر این است که برای حل مسئله جبران حرکت الگوریتم Sift نسبت به سایر الگوریتم‌ها بهتر عمل کرده و الگوریتم MinEigen بدترین عملکرد را داشته است.

**واژه‌های کلیدی:** ثبت تصاویر، شناسایی عوارض، تناظریابی، عوارض متحرک

## ۱- مقدمه

تشخیص خودروها در تصاویر ویدئویی به عنوان یک موضوع کلیدی در مباحث تحلیل ترافیک و بینایی کامپیوترا محسوب می‌گردد [۱]. اولین مرحله در سیستم‌های ردهیابی، آشکارسازی عوارض متحرک در تصاویر می‌باشد [۲]. در سال‌های اخیر، روش‌های گوناگونی به منظور استخراج و تشخیص عوارض متحرک با استفاده از دوربین‌های ناظری پیشنهاد شده است. فرض اصلی اکثر دوربین‌های ناظری، ثابت بودن دوربین در حین فیلمبرداری است [۳]. در چنین حالاتی، اگر دوربین ناظری 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].

تاکنون روش‌های مختلفی برای جبران حرکت دوربین ناظری انجام شده است. به عنوان مثال لیم سال‌های ۲۰۰۵ و ۲۰۱۳ در [۴، ۵] و همچنین چن سال در [۶] روش‌هایی ارائه کرده بودند که در آن با استفاده از تناظریابی و تفاضل بین دو فریم متوالی به مسئله جبران حرکت دوربین پرداخته بود.

در [۷] برای جبران حرکت دوربین از تناظر بین مجموعه‌های از ویژگی‌ها در دو فریم متوالی استفاده شده است.

در [۸] برای ردهیابی اشیاء توسط یک دوربین متحرک از روش جریان نوری استفاده شده است. در [۹] ردهیابی اشیاء متحرک بر اساس ویژگی با استفاده از یک دوربین دارای حرکات انتقالی، مورد بررسی قرار می‌گیرد. در [۱۰] نیز از ویژگی‌های گوشه بدست آمده با الگوریتم FAST برای جبران حرکت دوربین استفاده شد. در [۱۱] نیز بر اساس ویژگی گوشه به جبران حرکت دوربین و ردهیابی پرداخته‌اند. در [۱۲] 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 روزسانی می‌کنند. در [۱۳] 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.

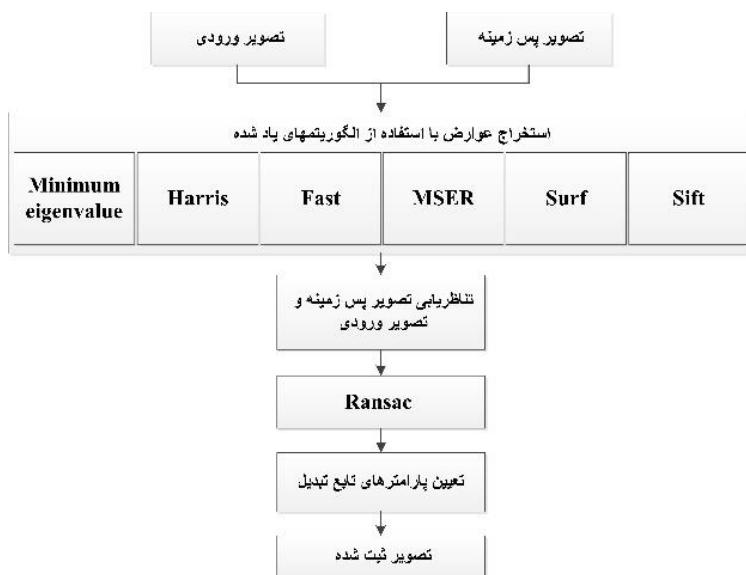
همان گونه که در اکثر روش‌های ذکر شده مشاهده می‌شود، برای حل مسئله جبران حرکت دوربین با استفاده از استخراج ویژگی، تناظریابی و تفاضل بین دو فریم متوالی  $t$  و  $t+1$  عمل شده است. اگر دوربین مورد استفاده خطای تاخیر احتمالی(Delay) (داشته باشد، دقت روش‌های یاد شده برای جبران حرکت تحت تاثیر این خطای قرار می‌گیرد). 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 Delay داشته باشیم هیچ تاثیری بر روی دقت جبران حرکت دوربین ندارد و کاملاً مستقل از این خطای عمل می‌کند. تنها خطایی که باقی می‌ماند تعییر در موقعیت مکانی پیکسل‌ها می‌باشد که با استفاده از ثبت تصاویر ورودی به تصویر پس‌زمینه (سیستم مختصات مرجع) این خطای هم جبران می‌شود و بهتر به شناسایی خودروها می‌توان پرداخت. بنابراین هدف اصلی در این تحقیق مقایسه ویژگی‌های مختلف ثبت تصاویر جهت حل مسئله جبران حرکت دوربین در روش تفاضل پس‌زمینه می‌باشد.

ساختار کلی مقاله در ادامه به این صورت است که در بخش دوم متداول‌تری نحوه جبران حرکت تشریح می‌گردد. در بخش سوم چگونگی تشخیص خودروها بررسی می‌شود. در بخش چهارم به نحوه پیاده‌سازی اشاره شده است. در بخش پنجم ارزیابی روش پیشنهادی صورت گرفته است. جمع‌بندی کلی در بخش ششم صورت گرفته است.

## ۲- 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 [۱۴].

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. در شکل ۱ با مقایسه ویژگی‌های مختلف برای ثبت تصاویر، هدف رسیدن به یک روش مناسب و کارآمد برای حل مسئله جبران حرکت می‌باشد

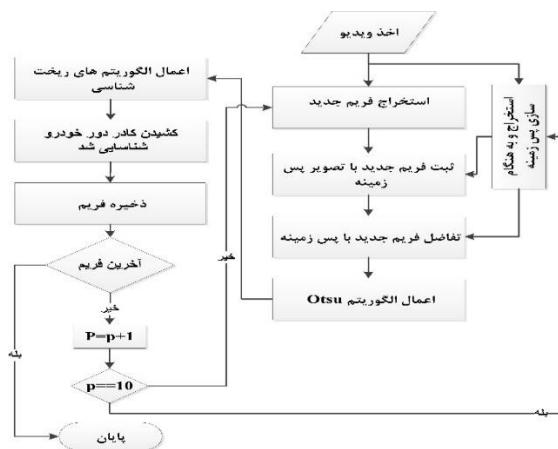


شکل ۱: فلوچارت مورد استفاده برای ثبت تصاویر

اولین مرحله برای ثبت دو تصویر، استخراج عوارض(ویژگی‌ها) می‌باشد. عوایضی که تاکنون برای برقراری ارتباط دو تصویر استفاده گردیده‌اند بسیار متنوع می‌باشند [۱۵]. در مقاله حاضر نیز شش الگوریتم برای استخراج عوارض در دو تصویر استفاده شده است. درادامه برای تناظریابی با مقایسه یک بردار توصیفگر از تصویر مرجع با تمام بردارهای توصیفگر نقاط کلیدی در تصویر هدف، به صورت ضرب داخلی دو بردار انجام می‌پذیرد [۶]. به علت اینکه بعضی از انطباق‌های اشتباه در مرحله تناظریابی تصاویر ممکن است رخ دهد، با استفاده از الگوریتم رنسک جفت تناظرهای اولیه بررسی شده و انطباق‌های غیر صحیح حذف می‌شوند [۱۶]. با توجه به موقعیت و شرایط فیلمبرداری که داریم امکان وجود تبدیل‌هایی چون تبدیل افاین و پروجکتیو... در بین دو تصویر مبنا و ورودی بالاست [۱۷]. در این تحقیق از تبدیل افاین استفاده شده است.

## ۳- تشخیص خودروها

برای تشخیص خودروها تصویر پس زمینه را از هریک از تصاویر ورودی کم می‌کنیم تا عوارض متحرک از تصویر استخراج شوند. در شکل ۲ فلوچارت پیشنهادی برای شناسایی خودروها نشان داده شده است.



شکل ۲: الگوریتم پیشنهادی برای شناسایی خودروها

برای استخراج خودروها، همان‌طور که در شکل (۲) نشان داده شده است، بعدازینکه 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. در این مقاله از روش فیلتر میانه برای استخراج تصویر پس‌زمینه استفاده شده 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.

بعدازین که عملیات ثبت روی فریمهای شناسایی عوارض متحرک، هر تصویر ورودی با تصویر پس‌زمینه تفاصل گرفته می‌شود.

$$Dif_n = |updated\ background_p - F_n| \quad (1)$$

تصویر پس‌زمینه به هنگام رسانی شده مرحله  $p$ ، و  $F_n$  تصویر هم مرجع شده  $n$  می‌باشد.

#### ۴- پیاده‌سازی

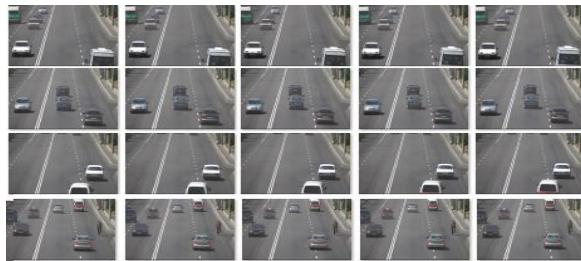
برای پیاده‌سازی مقایسات پیشنهادی از نرم‌افزار متلب استفاده شده است. داده‌های مورد استفاده برای تست روش پیشنهادی، از یک پایگاه داده مربوط به کشور اکراین استفاده شده است. مشخصات مربوط به داده مورد استفاده، در جدول (۱) تشریح شده است.

جدول ۱: داده‌های مورد استفاده برای ارزیابی سیستم پیشنهادی

نام داده	مکان	میزان لرزش دوربین	تراکم خودروها	زمان فیلم (دقیقه)	اندازه فریم (پیکسل)	فریم بر ثانیه	فرمت
Highway1	بکی از بزرگراه‌های کشور اکراین	متوسط	متوسط	۲۰	۳۶۰*۴۸۰	۲۵	MP4

ویژگی اصلی این 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 جعبه‌ابزار 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.



شکل ۳ : نمونه‌ای از فریم‌های استخراج شده

در این تحقیق برای استخراج خودروها در فریم‌های متوالی، از روش تفاضل پس‌زمینه استفاده شده است. بدین ترتیب که پس از استخراج تصویر پس‌زمینه، تمام فریم‌های ورودی را یک به یک با تصویر پس‌زمینه با استفاده از ویژگی‌های مختلف ثبت تصاویر به هم ثبت می‌کنیم. در هر یک از روش‌های ثبت، ویژگی‌های مورد نظر برای تصویر پس‌زمینه و هر فریم ورودی با الگوریتم‌های ذکر شده در شکل (۱) به دست آمده و سپس تناظریابی صورت گرفته و با استفاده از یکتابع تبدیل افاین می‌توان این تصاویر را ببروی هم ثبت کرد. تصویر پس‌زمینه از هریک از فریم‌های ورودی ثبت شده تفriق خواهد شد و حاصل این تفريق تصویری خواهد بود که در آن تنها عوارض غیرثابت (خودروها و ...) حضور دارند. شکل (۴) نمونه‌ای از تفاضل پس‌زمینه با فریم‌های متوالی برای استخراج اشیای متحرک موجود بعد از ثبت را نشان می‌دهد.

همانطور که در شکل ۴ نشان داده شده، نتایج مربوط به هریک از روش‌های ثبت در یک سطر به نمایش گذاشته شده است که ستون اول تفاضل پس‌زمینه با فریم ورودی می‌باشد. ستون دوم خودرو به صورت باينری و ستون سوم خودروی شناسايي شده در فریم اصلی را نشان می‌دهد.

#### ۵- ارزیابی

در این مقاله برای ارزیابی توانایی روش‌های مورد مقایسه در حذف حرکات اضافی دوربین نظارتی برای شناسایی دقیق‌تر عوارض متحرک، از متريک‌های مختلفی نظير FAR ، HR [۱۸] و MODP [۱۹] استفاده شده است.

$$HR = \frac{H}{H + M} \quad \text{رابطه (۲)}$$

$$FAR = \frac{F}{H + F} \quad \text{رابطه (۳)}$$

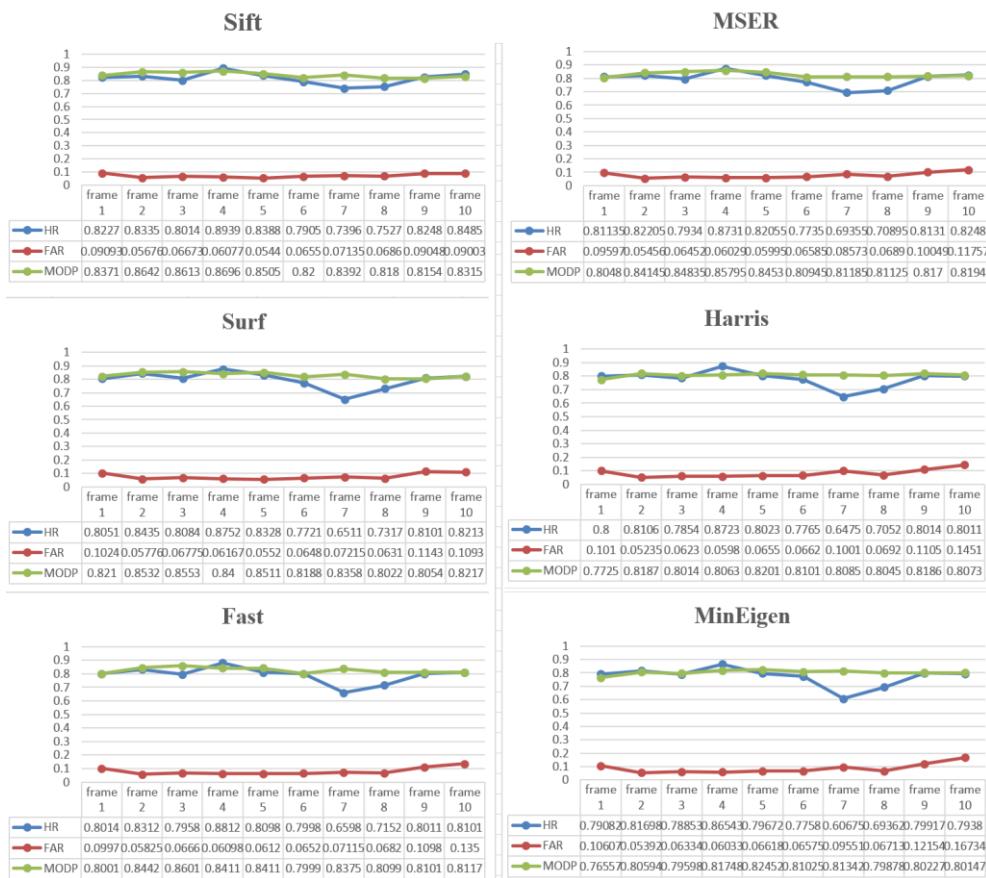
$$MODP = \frac{1}{N_{mapped}^t} \sum_{i=1}^{N_{mapped}^t} \left| \frac{G_i^{(t)} \cap D_i^{(t)}}{G_i^{(t)} \cup D_i^{(t)}} \right| \quad \text{رابطه (۴)}$$

در این روابط، H ، F و M به ترتیب نواحی درست، اشتباه و تشخیص نداده شده می‌باشند.  $G_i^{(t)}$  و  $D_i^{(t)}$  بیانگر عارضه واقعی و استخراج شده نام در فریم  $t$  می‌باشند. همچنین  $N_{mapped}^t$  مشخص کننده تعداد عارضه تصویر شده در فریم  $t$  می‌باشد. در ادامه برای بدست آوردن میزان دقت و قابلیت اعتماد در تشخیص خودروها از معیارهایی که در روابط ۳ تا ۵ بیان شده استفاده می‌شود. برای این کار به صورت اتفاقی ۱۰ فریم از داده را انتخاب می‌کنیم. خودروها را با استفاده از روش پیشنهادی استخراج می‌کنیم و از طرف دیگر پیکسل‌های مربوط به

موقعیت دقیق خودروها در این ۱۰ فریم به صورت دستی استخراج شده و به عنوان داده‌های واقعی با روش پیشنهادی مقایسه می‌شود. نتایج در قالب نمودارهای شکل ۵ نشان داد شده است.

روش	تفاضل فریم ورودی از پس زمینه	خودروهای شناسایی شده به صورت باینری	نشان دادن خودرو با یک مستطیل محاطی
SIFT			
Surf			
MSER			
Fast			
Harris			
MinEigen			

شکل ۴: ستون اول) تفاضل فریم‌های متوالی از تصویر پس زمینه. ستون دوم) خودروهای شناسایی شده به صورت باینری. ستون سوم) نشان دادن خودرو با یک مستطیل محاطی



شکل ۵ : نتایج بدست آمده برای ۱۰ فریم از ۳ مجموعه تصاویر ویدیویی برای ثبت تصاویر با استفاده از ۶ نوع ویژگی

به طور کلی با توجه به نمودارهای شکل ۵ هرچه میزان HR و MODP زیاد باشد نشان دهنده شناسایی بهتر و دقیقتر خودروها می‌باشد و بر عکس هر قدر FAR کمتر باشد نشان دهنده دقیق‌تر در شناسایی خودروها می‌باشد. در هر کدام از الگوریتم‌های ذکر شده، هر چقدر میزان HR و FAR بیشتر و MODP کمتر باشد به این معنی است که آن الگوریتم در مسئله جبران حرکت بهتر عمل کرده است. جدول زیر میانگین این معیارها را در ۱۰ فریم را نشان می‌دهد.

جدول ۲ : میانگین معیارهای ارزیابی برای ۵ فریم

Method	Sift	Surf	MSER	Fast	Harris	MinEigen
Average HR	<b>0.81464</b>	0.79513	0.79054	0.78023	0.793435	<b>0.77276167</b>
Average FAR	<b>0.071555</b>	0.076843	0.079608	0.083205	0.07738	<b>0.08671033</b>
Average MODP	<b>0.84068</b>	0.83045	0.82557	0.8068	0.82668	<b>0.8035675</b>

همانطور که از جدول ۲ پیداست میزان HR و MODP در الگوریتم sift بیشترین و الگوریتم MinEigen کمترین مقدار را دارند. همچنین FAR در الگوریتم Sift کمترین و در الگوریتم MinEigen بیشترین مقدار را دارد. به طور کلی این نتایج نشان بر این است که ویژگی سیفت برای حل مسئله جبران حرکت دوربین بهتر عمل کرده و ویژگی MinEigen عملکرد بدتری داشته است.

## ۶-نتیجه‌گیری

در این تحقیق مقایسه ویژگی‌های مختلف ثبت تصاویر صورت گرفته و روشی مناسب و کارآمد برای جبران حرکت اضافی دوربین‌های نظارتی با استفاده از ثبت تصاویر متوالی با پس‌زمینه در شناسایی خودروها، ارائه شده است. برای این منظور بعد از استخراج فریم‌ها و پیدا کردن تصویر پس‌زمینه، باید تمام فریم‌های ورودی با تصویر پس‌زمینه براساس الگوریتم‌های پیشنهادی، ثبت شوند. سپس با استفاده از روش تفربیق پس‌زمینه به شناسایی خودروها پرداخته شده است. این امر باعث حذف خطای Delay در حین تصویر برداری می‌شود. برای ارزیابی نتایج، روش پیشنهادی بر روی یک داده با شش روش پیاده‌سازی شده است. ویژگی اصلی این مجموعه داده‌ها لرزش‌های مکرر دوربین در حین فیلم‌برداری و متفاوت بودن تراکم خودروهای موجود در هر مجموعه داده می‌باشد. از متريک‌های MODP, HR, FAR برای مقایسه روش‌های موجود و تعیین بهترین الگوریتم برای جبران حرکت استفاده شده است. با توجه به نتایج به دست آمده الگوریتم سيفت بهترین عملکرد و همچنین MinEigen بدترین عملکرد را در حل مسئله جبران حرکت داشته‌اند.

## مراجع

- [1] A. Karamiani and N. Farajzadeh, "Optimal feature points for tracking multiple moving objects in active camera model," *Multimedia Tools and Applications*, pp. ۱۹-۱, .۲۰۱۵
- [2] R. Poppe, "Vision-based human motion analysis: An overview," *Computer vision and image understanding*, vol. ۱۰۸, pp. ۱۸-۴, .۲۰۰۷
- [3] M. Piccardi, "Background subtraction techniques: a review," in *Systems, man and cybernetics, ۲۰۰۴ IEEE international conference on*, ۲۰۰۴, pp. ۳۱۰۴-۳۰۹۹
- [4] J. S. Lim and W. H. Kim, "Detection and tracking multiple pedestrians from a moving camera," in *Advances in Visual Computing*, ed: Springer, ۲۰۰۵, pp. ۵۳۴-۵۲۷
- [5] J. Lim and W. Kim, "Detecting and tracking of multiple pedestrians using motion, color information and the AdaBoost algorithm," *Multimedia tools and applications*, vol. ۶۵, pp. ۱۷۹-۱۶۱, .۲۰۱۳
- [6] 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), ۲۰۱۲IEEE International Conference on*, ۲۰۱۲, pp. ۶۹۷-۶۹۶
- [7] 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*, ۲۰۰۴, pp. ۹۸۷-۹۸۰
- [8] 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, ۲۰۰۳.(IROS ۲۰۰۳). Proceedings. ۲۰۰۳IEEE/RSJ International Conference on*, ۲۰۰۳, pp. ۱۳۱۳-۱۳۰۸ .
- [9] 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 ۲۰۰۲*, ed: Springer, ۲۰۰۲, pp. ۱۱۴۴-۱۱۳۷
- [10] M. Siam, R. ElSayed, and M. ElHelw, "On-board multiple target detection and tracking on camera-equipped aerial vehicles," in *Robotics and Biomimetics (ROBIO), ۲۰۱۲IEEE International Conference on*, ۲۰۱۲, pp. ۲۴۰۵-۲۳۹۹

- [11] 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. 207-215, 2014
- [12] 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. 552-544
- [13] 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. 155-145, 2009
- [14] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, pp. 100-177, 2003
- [15] L. Xiaoming, Z. Lian, and H. Zhanyi, "SIFT based automatic registration of remotely-sensed imagery," *Journal of remote sensing*, vol. 10, pp. 892-885.
- [16] 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. 384-395, 1981
- [17] J. L. Mundy and A. Zisserman, *Geometric invariance in computer vision* vol. 92: MIT press Cambridge, MA, 1992
- [18] 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. 617-628
- [19] 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