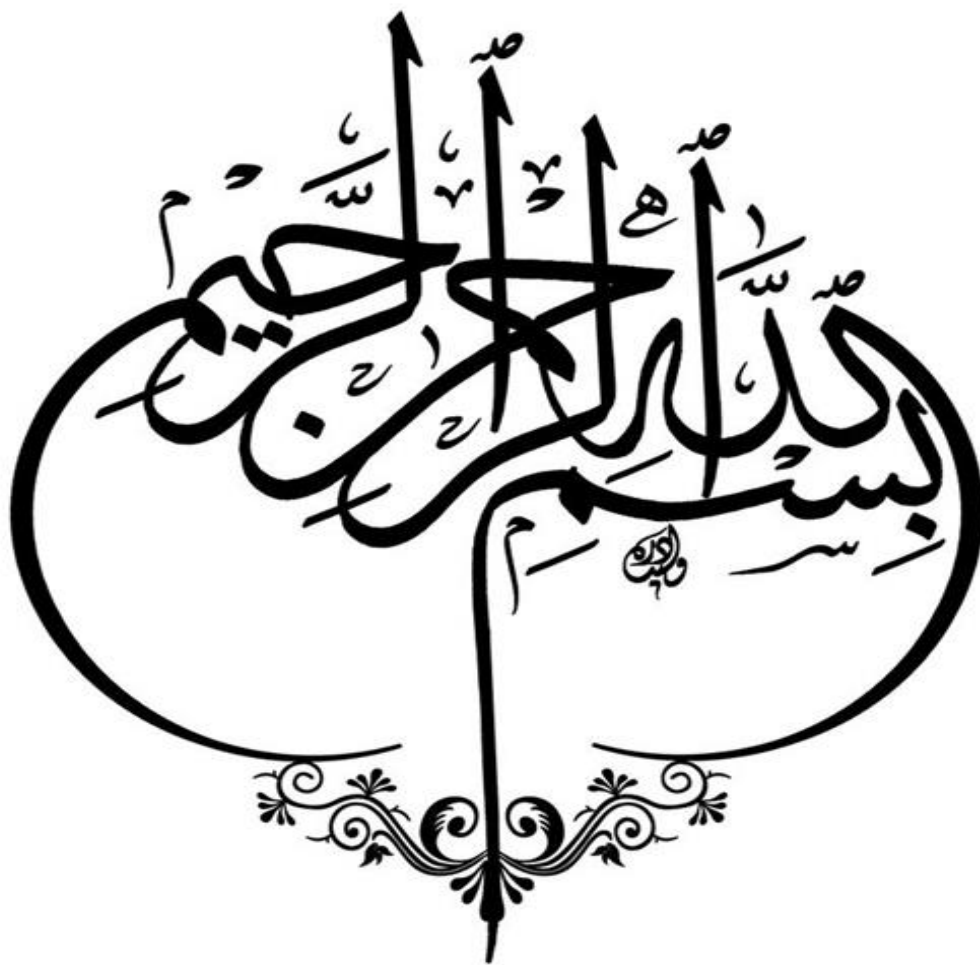
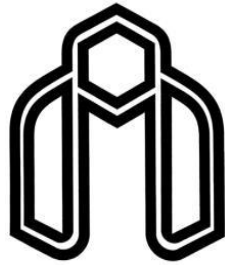






دانشگاه صنعتی شاهرود

دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد مهندسی هوش مصنوعی

بهبود عملکرد پایدارسازی تصاویر ویدئو بر اساس محتوای صحنه

نگارنده: سیاوش ریاحی ثالث

استاد راهنما

آقای دکتر حمید حسن پور

استاد مشاور

آقای دکتر محسن بیگلری

خرداد ۱۴۰۰

تقدیم به خانواده عزیزم

سرمایه‌هایی که با نهایت دلسوزی در تمامی مراحل زندگی، عامل رشد و بالندگی‌م بودند و هستند،
دل‌بستگی‌های ارزشمندم که همواره در کنارم بوده‌اند و حضور سبزشان در زندگی‌م آرزوی
همیشگی‌م است.

سپاس‌گزاری

از استاد بی‌بدیل و گران‌قدرم جناب آقای دکتر **حمید حسن‌پور** که درس‌های گران‌بها از ایشان آموختم و در انجام این پایان‌نامه از هیچ کمکی فروگذار نکردند، بی‌اندازه سپاس‌گزار هستم.

از جناب آقای دکتر **محسن بیگلری** که در تمامی مراحل انجام این پایان‌نامه همراه و در کنار بنده بودند نیز سپاس‌گزارم.

سیاوش ریاحی ثالث

خرداد ۱۴۰۰

تعهدنامه

اینجانب **سیاوش ریاحی ثالث** دانشجوی کارشناسی ارشد رشته **هوش مصنوعی** دانشکده **کامپیوتر و فناوری اطلاعات** دانشگاه شاهرود، نویسنده پایان نامه با عنوان **بهبود عملکرد پایدارسازی تصاویر ویدئو بر اساس محتوای صحنه**، تحت راهنمایی **آقای دکتر حمید حسن پور** متعهد می شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های دیگر پژوهش گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تاکنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام "دانشگاه صنعتی شاهرود" یا "Shahrood University of Technology" به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده اند، در مقالات مستخرج از پایان نامه رعایت می گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است) اصل رازداری و اصول اخلاق انسانی رعایت شده است.

سیاوش ریاحی ثالث

خرداد ۱۴۰۰

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان نامه بدون ذکر منبع مجاز نمی باشد.

چکیده:

تصاویر ثبت شده‌ی غیرحرفه‌ای معمولاً از لرزش و تکان‌های ناخواسته حس‌گر دوربین رنج می‌برند و معمولاً حرکت متزلزل یک دوربین دستی بی‌ثبات یا یک دوربین نصب شده روی خودرو، باعث به وجود آمدن نویز در فیلم ضبط شده می‌شود که این اتفاق، احساس آزاردهنده‌ای را برای بیننده به وجود می‌آورد. امروزه به‌منظور پایدارسازی تصاویر ویدیوی از روش‌های مختلفی استفاده می‌شود و افزایش میزان دقت در پایدارسازی تصاویر ویدیوی و همچنین کاهش میزان خطای مدل یکی از مهم‌ترین چالش‌ها در این حوزه می‌باشد، به همین دلیل در این پژوهش نیز از یک روش ترکیبی به‌منظور پایدارسازی فریم‌های ویدیو استفاده شده است.

در این مدل ابتدا فرایند تشخیص محتوای صحنه با استفاده از روش *Grow Cut* و الگوریتم زرنیک انجام شده است و در ادامه به‌منظور پایدارسازی تصویر از یک روش تطبیق بلوکی^۱ استفاده شده است، درواقع فرایند پایدارسازی تصویر در روش پیشنهادی شامل دو بخش تشخیص محتوای تصویر و پایدارسازی تصویر می‌باشد. در بخش اول فرایند تشخیص محتوای تصویر انجام می‌شود، درواقع انجام فرایند تشخیص محتوای تصویر شامل دو بخش، انتقال فریم به الگوریتم *Grow Cut* و قطعه‌بندی و سپس استخراج ویژگی و تشخیص تغییر محتوا با استفاده از الگوریتم زرنیک می‌باشد. در بخش دوم: فرایند پایدارسازی تصویر انجام می‌شود، در این قسمت به‌منظور پایدارسازی تصویر از روش تطبیق بلوکی استفاده شده است.

در ادامه و به‌منظور ارزیابی مدل پیشنهادی بر استفاده از ویدیو و سنجش بصری نتایج تاکید شده است اما همچنین از پارامترهای نمودار جابجایی سراسری (GMV^2)، میانگین و واریانس حاصل از مقادیر این پارامتر، میانگین مربعات جابجایی استفاده شده است. نتایج این پژوهش نشان‌دهنده عملکرد مطلوب روش ارایه شده در مقایسه با روش‌های متداول و همچنین روش پایه می‌باشد. مقادیر بدست آمده برای میانگین جابجایی در محور X برابر ۰.۴۵ و برای محور Y برابر ۰.۴۱۶ بوده و مقدار واریانس در محور X برابر ۰.۰۵۱ و در محور Y برابر ۰.۰۴۵ بوده است. همچنین میانگین مربعات

¹ Block Matching

² Global Motion Vector

جابجایی ویدیوی ورودی و روش پیشنهادی برابر ۰.۳۲۸ در راستای محور X و مقدار ۰.۱۳۳ در راستای محور Y بوده است و بهبود نتایج حاصل از روش پیشنهادی را نشان می‌دهد.

واژه‌های کلیدی: سیستم پایدارسازی تصویر، روش تطبیق بلوک، استخراج ویژگی، الگوریتم زرنیک، نمودار حرکت سراسری.

فهرست مطالب

فصل اول : کلیات تحقیق..... ۱

- ۱-۱- مقدمه ۲
- ۲-۱- کاربرد و چالش‌ها ۲
- ۳-۱- روش‌ها ۳
- ۴-۱- ضرورت انجام تحقیق ۴
- ۵-۱- روش و نوآوری تحقیق ۴
- ۶-۱- جمع‌بندی و ساختار پایان‌نامه ۶

فصل دوم: مروری بر کارهای پیشین..... ۷

- ۱-۲- مقدمه ۸
- ۲-۲- نويز ۸
- ۳-۲- دلایل ایجاد نويز ۹
- ۱-۳-۲- مقدمه‌ای بر تاري ۱۰
- ۲-۳-۲- لرزش دوربين در چه مواقعی موجب تاري می‌شود؟ ۱۲
- ۳-۳-۲- تاري ۱۴
- ۱-۳-۳-۲- تاري در برابر زوم کردن ۱۴
- ۲-۳-۳-۲- انواع فیلترها ۱۴
- ۳-۳-۳-۲- کاربردهای حذف تاري ۱۵
- ۴-۲- الگوریتم‌های کیفیت تصویر ۱۶
- ۱-۴-۲- روش‌های عینی ارزیابی کیفیت تصویر ۱۷
- ۵-۲- برخی از روش‌های حذف نويز ۱۷
- ۱-۵-۲- فیلتر وینر ۱۷
- ۲-۵-۲- فیلتر میانه ۱۸
- ۳-۵-۲- فیلتر گوسی ۱۹
- ۴-۵-۲- فیلتر SIFT ۱۹
- ۵-۵-۲- کاهش نويز در تصاویر با استفاده از روش میانگین‌گیری ۲۰
- ۶-۲- سوابق کار ۲۰
- ۷-۲- جمع‌بندی ۳۰

۳- فصل سوم: روش پیشنهادی..... ۳۱

- ۱-۳- مقدمه ۳۲

۳-۲-۳	روش پیشنهادی	۳۲
۳-۲-۱	انجام فرایند تشخیص محتوای صحنه	۳۴
۳-۲-۱-۱	نحوه پردازش تصاویر با استفاده از الگوریتم Grow Cut	۳۴
۳-۲-۱-۲	انجام فرایند استخراج ویژگی با استفاده از الگوریتم زرنیک	۳۶
۳-۲-۲	فرایند پایدارسازی تصویر	۳۷
۳-۳	جمع‌بندی	۳۹

۴- فصل چهارم: پیاده سازی مدل پیشنهادی

۴-۱	مقدمه	۴۲
۴-۲	پیاده‌سازی روش پیشنهادی	۴۲
۴-۲-۱	داده‌های مدل پیشنهادی	۴۲
۴-۲-۲	پارامترهای ارزیابی مدل پیشنهادی	۴۳
۴-۳	نتایج مدل پیشنهادی	۴۴
۴-۳-۱	ارزیابی نتایج مدل پیشنهادی	۴۵
۴-۳-۲	مقایسه‌ی نتایج مدل پیشنهادی با مقاله‌ی پایه	۴۷
۴-۴	نتیجه‌گیری	۵۲

۵- فصل پنجم: نتیجه‌گیری و پیشنهادها

۵-۱	نتیجه‌گیری	۵۴
۵-۲	کار آینده	۵۵

مراجع

فهرست اشکال

- شکل ۱-۱: نتیجه تصاویر پایدار شده. ردیف بالا تصاویر اصلی و ردیف پایین تصاویر پایدار شده [۱۶]..... ۵
- شکل ۱-۲: یک تصویر خاکستری..... ۸
- شکل ۲-۲: نمودار پایه گام‌های اساسی پردازش تصویر..... ۹
- شکل ۳-۲: انواع مختلف تاری در تصویر..... ۱۲
- شکل ۴-۲: یک تصویر ساده که در سطحی بزرگ تار و بلوره شده باشد..... ۱۳
- شکل ۵-۲: مقایسه میان تصویر اصلی و تصویر تار شده پس از اعمال ماسک..... ۱۵
- شکل ۶-۲: حذف تاری در تصاویر میکروسکوپی نجومی..... ۱۶
- شکل ۷-۲: فلوچارت SIFT..... ۲۰
- شکل ۸-۲: تخمین Tx از مدل حرکت عمدی دوربین با استفاده از فیلتر پایین گذر علی..... ۲۳
- شکل ۱-۳: روش پایدار سازی تصویر در مدل پیشنهادی..... ۳۳
- شکل ۲-۳: فلوچارت روش پیشنهادی..... ۳۴
- شکل ۳-۳: مراحل انجام فرایند تقسیم تعاملی در الگوریتم Grow Cut..... ۳۶
- شکل ۴-۳: مراحل پایدار سازی تصویر در مدل پیشنهادی..... ۳۷
- شکل ۵-۳: هیستوگرام زیر تصاویرهای انتخابی..... ۳۸
- شکل ۶-۳: نمونه ای از فرایند پایدار سازی تصویر..... ۳۹
- شکل ۱-۴: میزان جابجایی تصویر در محور X بر اساس ارقام GMV در مدل پیشنهادی..... ۴۶
- شکل ۲-۴: مقایسه‌ی خروجی مدل پیشنهادی با فریم‌های ورودی..... ۴۷
- شکل ۳-۴: میزان جابجایی تصویر در محور X بر اساس ارقام GMV در مقاله‌ی پایه..... ۴۹
- شکل ۴-۴: مقایسه‌ی خروجی مدل مقاله‌ی پایه با فریم‌های ورودی..... ۵۰
- شکل ۵-۴: مقایسه‌ی خروجی مدل پیشنهادی با خروجی مقاله‌ی پایه..... ۵۱

فهرست جداول

جدول ۱-۲: قوانین سیستم کنترل فازی پیشنهادی	۲۵
جدول ۱-۴: میزان واریانس مدل پیشنهادی با واریانس فریم‌های ورودی	۴۵
جدول ۲-۴: محاسبه‌ی میانگین جابجایی تصویر در محور X و Y بر اساس ارقام GMV	۴۵
جدول ۳-۴: مقایسه‌ی میزان واریانس مدل مقاله‌ی پایه [۸۰]، با فریم‌های ورودی	۴۸
جدول ۴-۴: مقایسه‌ی میانگین مربعات جابجایی مدل پیشنهادی با فریم‌های ورودی	۴۸
جدول ۵-۴: میانگین میزان جابجایی تصویر در محور X و Y ورودی، مقاله [۸۰] و روش پیشنهادی	۴۹

فصل ۱ – کلیات تحقیق

۱-۱- مقدمه

تصاویر ثبت شده توسط افرادی که هنگام حرکت اقدام به تصویربرداری می‌نمایند و یا دوربین نصب شده بر روی هواپیماهای بدون سرنشین^۱ دارای لرزش‌های ناخواسته‌ای هستند که سبب کاهش کیفیت آن‌ها می‌شود. در دهه‌های اخیر و با افزایش تعداد دوربین‌های قابل حمل و قیمت ارزان آن‌ها، استفاده از این دوربین‌ها و میزان تصاویر و ویدئوهای ثبت شده توسط آن‌ها روزبه‌روز افزایش یافته است [۱]. به‌طور مثال هواپیماهای بدون سرنشین که در عملیات نجات هوایی، مدیریت بیماری‌ها، تحویل کالا، تصویربرداری هوایی، اقدامات نظارتی و امنیتی کاربردهای فراوان دارند [۲]. پایداری تصاویر پردازشی مهم برای استفاده از این تصاویر محسوب می‌گردد.

۱-۲- کاربرد و چالش‌ها

در کنار این مزیت‌ها تصاویر ثبت شده توسط این‌گونه دوربین‌ها اغلب دچار مشکلاتی مانند نویز متأثر از لرزش‌های ناخواسته دست توسط افراد غیرحرفه‌ای [۳]، لرزش‌های موتور و یا جریان هوا در هنگام پرواز می‌باشند [۱]. حذف این لرزش‌ها و به دست آوردن فریم‌های باثبات بیشتر به‌عنوان پیش‌پردازشی مهم برای کاربردهای دیگری مانند شناسایی و تعقیب اجسام، شناسایی و تشخیص چهره در تصاویر ثبت شده موردنیاز است [۴]. جهت دستیابی به این هدف چند رویکرد کلی مطرح شده است [۵]–[۷] که روش نرم‌افزاری آن پایداری دیجیتالی تصویر (DIS)^۲ نام دارد. این روش از تکنیک‌های پردازش تصویر برای کنترل پایداری تصاویر بهره می‌برد؛ و متفاوت از آنچه در روش‌های دیگر انجام می‌شود، در این روش تصویر پس از آن‌که به حس‌گر دوربین رسید و مورد پردازش قرار می‌گیرد. روش‌های مبتنی بر تکنیک‌های پردازش

¹ Unmanned Aerial Vehicle

² Digital Image Stabilization

تصویر نسبت به روش‌های دیگر در این حیطه از لحاظ اقتصادی به صرفه بوده و قابلیت پس‌پردازش را برای داده‌ها فراهم می‌آورد.

از چالش‌های مربوط به حوزه پایدارسازی تصاویر ویدیویی می‌توان به تنوع ویدیوهای ورودی اشاره کرد، این موضوع منجر به تعدد نمونه‌های ورودی شده و در هر روش پیشنهادی جدید نیاز به بررسی نوع ویدیو یا حتی تولید ویدیو مدنظر محقق می‌شود.

کمبود پارامترهای آماری و ریاضی جهت ارزیابی کیفیت نتایج حاصل از روش پیشنهادی و میزان بهبود خروجی از دیگر چالش‌های پیش رو در این حوزه می‌باشد. در نتیجه اکثراً این روش‌ها نیازمند قضاوت بصری انسانی جهت میزان بهبود حاصل شده می‌باشند.

۱-۳- روش‌ها

به‌طور کلی تکنیک‌های پایدارسازی تصاویر به سه روش تقسیم می‌شوند: [۵]

پایدارسازی اپتیکال تصاویر^۱: در این روش تصویر قبل از رسیدن به سنسور دوربین دستکاری می‌شود.

پایدارسازی مکانیکی تصاویر^۲: این روش شامل کل دوربین می‌شود که با استفاده از سنسوری مانند ژيروسکوپ لرزش‌ها تشخیص داده شده و به صورت مکانیکی جبران می‌شوند.

پایدارسازی دیجیتالی تصاویر^۳: این روش از تکنیک‌های پردازش تصویر برای کنترل پایداری تصاویر بهره می‌برد. هر دو روش اول و دوم از سخت افزار برای پایدارسازی استفاده می‌کنند اما در پایدارسازی دیجیتالی، تصویر پس از آن‌که به سنسور دوربین رسید و ثبت شد مورد پردازش قرار می‌گیرد. روش‌های دیجیتالی نسبت به روش‌های مکانیکی از لحاظ اقتصادی بصره بوده و همچنین موجب افزایش وزن سنسورهای ثبت تصاویر نیز نمی‌شوند.

روش‌های مبتنی بر تکنیک‌های پردازش تصویر عموماً شامل سه مرحله هستند

^۱ Optical Image Stabilization (OIS)

^۲ Mechanical Image Stabilization (MIS)

^۳ Digital Image Stabilization (DIS)

تخمین حرکت^۱: این مرحله به منظور تخمینی از میزان جابجایی سنسور در هر فریم نسبت به فریم بعدی صورت می‌گیرد.

جبران حرکت^۲: در این گام میزان جابجایی بوجود آمده جبران می‌شود تا بتوان حرکات ناخواسته سنسور را مرتفع کرد.

ترکیب تصویر^۳: ساخت فریم جدید و بدون لرزش با توجه به مراحل قبلی [۶]

۴-۱- ضرورت انجام تحقیق

همان‌طور که بیان شد، تصاویر ثبت شده غیرحرفه‌ای معمولاً از لرزش و تکان‌های ناخواسته حس‌گر دوربین رنج می‌برند و نیازمند پیش‌پردازش جهت کاربردهای آتی می‌باشند. استفاده از تصاویر ویدئوها امروزه بیش از پیش صورت می‌گیرد و کاربردهای گوناگون و با اهداف مختلف مانند بهره‌وری نظامی، اجتماعی، تجاری و آموزشی موجب افزایش ثبت تصاویر می‌شود. رشد روزافزون فناوری و کوچک‌تر شدن تجهیزات تصویربرداری امکان حمل راحت‌تر و نصب آن‌ها را بر روی تلفن‌های همراه، ربات‌ها و یا هواپیماهای بدون سرنشین را فراهم کرده است. روش‌های سخت‌افزاری عموماً هزینه‌بر هستند و نیازمند نصب تجهیزاتی بر روی دوربین‌ها می‌باشند. از این‌رو روش‌های نرم‌افزاری ارائه شده برای پایدارسازی تصاویر مورد توجه واقع شده‌اند. [۵] و [۶]

۵-۱- روش و نوآوری تحقیق

اقدامات کلی در انجام پایدارسازی تصاویر نیازمند این است که از تعدادی فریم متوالی، نواحی مشترک استخراج شود سپس با در نظر گرفتن این نواحی و حذف بخش‌های دیگر فریم، تصویر حاصل پایدارسازی شود؛ اما این عمل معمولاً موجب محدود شدن فریم به ROI^4 می‌شود. ROI به بخشی از تصویر گفته

¹ Motion estimation

² Motion compensation

³ Image composition

⁴ Region of Interest

می‌شود که شامل محتوایی است که می‌خواهیم عملیاتی روی آن انجام دهیم. یافتن این نواحی از تصاویر موجب کاهش محاسبات می‌شود. می‌توان ROI را با شکل‌های زیاد و با کمک توابع آن تولید کرد. این نواحی به‌دست آمده معمولاً دارای شباهت‌هایی هستند که آن‌ها را از نواحی دیگر متمایز می‌کند. یافتن نواحی مشترک از فریم‌ها، جهت پایداری تصاویر کمک می‌کند.



شکل ۱-۱ نتیجه تصاویر پایدار شده. ردیف بالا تصاویر اصلی و ردیف پایین تصاویر پایدار شده [۱۶]

روش‌های موجود از ویژگی‌های سطح پائین برای انجام این هدف استفاده می‌کنند. یکی از مسائل حائز اهمیت میزان فضای از دست رفته در هر فریم می‌باشد. هرچه بازه فریم‌ها بیشتر باشد، میزان کمتری از تصویر باقی می‌ماند. با محدود کردن تعداد فریم‌های در نظر گرفته شده برای پایداری در هر مرحله، می‌توان این نواحی اشتراک را کاهش داد. رویکرد مورد نظر برای غلبه بر این مشکل، پایداری تصاویر بر اساس محتوای فریم‌هاست. ابتدا محتوای فریم‌ها مشخص می‌شود و برای هر موضوع بر روی فریم‌های مربوط به آن، عملیات پایداری صورت گیرد و سپس برای بازه دیگری از فریم‌ها و در صورت تغییر موضوع، این عملیات مجدداً صورت می‌پذیرد. به‌طور مثال اگر f_t فریم در زمان t باشد و محتوای تصویر تا f_{t+n} یکسان باشد برای n فریم این بازه نیاز به یک مرحله انجام عملیات و در صورت تغییر محتوا از f_{t+n} تا f_{t+m} نیاز هست که مجدداً برای بازه فریم‌های n تا m این عملیات صورت پذیرد.

درواقع در این پژوهش ابتدا فرایند تشخیص محتوای تصویر با استفاده از روش Grow Cut و الگوریتم زرنیک انجام می‌شود و همچنین در ادامه به منظور پایداری تصویر از یک روش تطبیق بلوکی استفاده شده است.

۱-۶- جمع‌بندی و ساختار پایان‌نامه

همان‌طور که مشاهده شد در این فصل ابتدا به بیان مسئله و مشکلات موجود در زمینه حذف نویز در تصاویر پرداخته شده است و همان‌طور که گفته شد در این پژوهش به منظور حذف نویز و پایداری تصاویر در قسمت تشخیص محتوای تصویر از الگوریتم زرنیک استفاده شده است و به منظور پایداری تصویر از روش تطبیق بلوک استفاده شده است، در ادامه در فصل دوم به توضیح کارهای انجام شده در حوزه حذف نویز و پایداری تصویر می‌پردازیم و بررسی خواهیم کرد که بهترین روش در زمینه حذف نویز کدام روش می‌باشد. در فصل سوم به توضیح کامل مدل پیشنهادی خواهیم پرداخت و در فصل چهارم به بررسی نتایج مدل پیشنهادی پرداخته و خروجی مدل پیشنهادی را مورد ارزیابی قرار می‌دهیم و سپس به مقایسه‌ی خروجی مدل پیشنهادی با مدل‌های دیگر می‌پردازیم؛ و در نهایت در فصل پنجم ابتدا مروری بر کارهای انجام شده در این پژوهش خواهیم پرداخت و در ادامه به بررسی نقاط قوت و ضعف مدل پیشنهادی می‌پردازیم.

فصل ۲- مروری بر کارهای پیشین

۲-۱- مقدمه

با فراگیری استفاده از دوربین‌ها و افزایش روزافزون تصاویر ثبت شده و در کنار آن عدم وجود تخصص کافی و یا سخت‌افزار لازم جهت ثبت تصاویر ویدئویی باکیفیت، طی سالیان متمادی روش‌های زیادی در جهت بهبود این تصاویر و در زمینه‌های مختلف برای رسیدن به این هدف ارائه شده است. این روش‌ها با توجه به کاربردی که خواهند داشت، نوع و منشأ عدم کیفیت تصویر، نحوه استفاده و سرعت عمل لازم جهت پردازش تصاویر دسته‌بندی‌های بسیاری دارند.

۲-۲- نویز

نویزها باعث می‌شوند که سطح جزئیات ثبت شده در عکس دیجیتالی یا فیلمی کاهش یابد، بنابراین کاهش این نویزها سبب می‌شود که کیفیت عکس نهایی یا خروجی چاپی افزایش یابد. مشکلی که وجود دارد این است که تکنیک‌های رایج برای کاهش یا حذف نویز تصاویر به همان اندازه منجر به صاف شدن تصویر می‌شود. صاف شدن تصویر در عکس‌هایی که از موضوعاتی مانند آسمان یا سطح آب گرفته شده‌اند مشکل چندانی ایجاد نمی‌کند اما مثلاً در تصویر یک منظره از درختان حتی تلاش‌های محتاطانه برای کاهش نویز موجب صدمه دیدن جزئیات شاخ و برگ‌های درختان موجود در عکس می‌شود. [۸]

نویزها در عکاسی دیجیتال و اسکن انواعی دارند که دو مورد آن‌ها از بقیه شایع‌تر هستند نویز تصادفی^۱ و نویز با الگوی ثابت^۲. نویزهای تصادفی غالباً در ایزوهای بالا دیده می‌شوند و همان‌گونه که از اسم‌شان مشخص است، محل ثابتی ندارند؛ یعنی اگر دو یا چند عکس با شرایط یکسان و با ایزو بالا تهیه شود، محل نویزها در عکس‌ها مختلف متفاوت است؛ اما نویزهای ثابت غالباً در نوردهی‌های طولانی و موارد دیگری که با گرم شدن حس‌گر همراه است، دیده می‌شوند.

حذف نویزهای ثابت راحت‌تر است چون محل آن‌ها ثابت است و اگر یک فریم سیاه (با بستن درپوش لنز) بگیریم و تفاضل آن را با عکس اصلی تهیه کنیم، نویزها حذف خواهند شد. [۹] کاری که برخی دوربین‌ها

¹ Random Noise

² Fixed Pattern Noise

به‌طور خودکار نیز انجام می‌دهند؛ اما چالش اصلی در حذف نویزهای اتفاقی است؛ چه در دوربین، اسکنر و چه در نرم‌افزار. غالباً این روش‌ها منجر به کاهش جزئیات عکس و یا ایجاد آثار جانبی، بخصوص در لبه‌های اجسام می‌گردند. درواقع کاربر باید بین این مشکلات و نویز، تعادلی برقرار نماید و تصویر نهایی، حاصل این تصمیم کاربر خواهد بود.

۲-۳- دلایل ایجاد نویز

نویز معمولاً در تصویر وجود دارد به این معنی که هرگز تصویر ثبت شده مانند واقعیت نیست، اما میزان آن می‌تواند آن‌قدر کم باشد که بتوان گفت وجود ندارد و یا آن‌قدر زیاد باشد که نتوان سوژه‌ی تصویر را تشخیص داد. نویز به دلایل مختلفی ایجاد می‌شود و می‌تواند ناشی از یک قسمت و یا مجموع چندین قسمت باشد.

در بخش ثبت تصویر در دوربین‌های دیجیتال، بازماندن طولانی شاتر منجر به ایجاد نقاط سفید در بخش سیاه و برعکس می‌شود که ناشی از کاهش جریان دیودهای نوری است.

در حساسیت (ISO) های بالاتر، سیگنال‌های به‌دست آمده تقویت می‌شوند تا نتیجه‌ای همانند حساسیت‌های پایین به‌دست آید و همین تقویت سیگنال‌ها باعث افزایش نویز می‌شود. اندازه‌ی حس‌گر هم نقشی به‌سزا در میزان نویز دارد، حس‌گرهای بزرگ‌تر، بخش بیشتری از نوری که از لنز عبور می‌کند را استفاده می‌کنند و حساسیت و سرعت شاتر کمتری لازم خواهند داشت. یکی از دلایلی که عکاس‌های حرفه‌ای، بزرگی و سنگینی دوربین‌های خود را تحمل می‌کنند همین بزرگی حس‌گرهایشان است (چند سال اخیر دوربین‌های کوچکی نیز با حس‌گرهای بزرگ ساخته شده‌اند)، پس نباید از حس‌گرهای کوچک مثل موبایل‌ها و دوربین‌های فشرده انتظار زیادی داشت. همچنین، نوری که به هر نقطه (پیکسل) از حس‌گر می‌رسد، با حاصل تقسیم مساحت حس‌گر بر تعداد نقاط آن رابطه مستقیم دارد، درنتیجه در دو حس‌گر با اندازه یکسان با تعداد نقطه‌های متفاوت، سنسوری که نقطه‌های کمتری دارد (پیکسل کمتر)، نور بیشتری برای هر یک از نقطه‌ها فراهم کرده و نویز کمتری دارد، پس تعداد پیکسل بیشتر بر خلاف دید عموم، علاوه بر

اینکه منجر به بهبود کیفیت نمی‌شود، نور کمتری برای هر نقطه فراهم کرده و می‌تواند منجر به ایجاد نویز بیشتری گردد.

به همین دلیل، دوربین‌های سریع که برای رخدادهای ورزشی استفاده می‌شوند تعداد نقاط کمتری در حس‌گرهای خود نسبت به دوربین‌هایی که در آتلیه‌ها استفاده می‌شوند دارند تا بتوانند در سرعت‌های بالاتر نویز کمتری ایجاد کنند. دما نیز می‌تواند بر میزان نویز حس‌گر اثر بگذارد، افزایش دما منجر به افزایش نویز می‌شود، در نتیجه دوربین‌ها در تابستان بیشتر از زمستان نویز ثبت می‌کنند. خطا در تبدیل اطلاعات آنالوگ به دیجیتال، سرعت پایین پردازش اطلاعات، روش‌های تبدیل و ذخیره‌سازی، پایداری ولتاژ باتری‌ها و... نیز می‌توانند منجر به ایجاد نویز شوند.

همچنین ممکن است نقطه‌هایی از حس‌گر به دلیل آسیب‌دیدگی سوخته باشند و نقطه‌های سیاهی را در تصویر ایجاد کنند که با توجه به تکرار شدن در همه‌ی تصویرها، از نویز تصادفی قابل تشخیص هستند. در دوربین‌های آنالوگ، بازماندن شاتر هیچ نویزی ایجاد نمی‌کند و تنها بخش تولیدکننده‌ی نویز، فیلم آن است که با توجه به گونه، تاریخ تولید، شرایط نگهداری، اندازه و حساسیت مواد حساس به نور (هالیدهای نقره) می‌تواند مقدار و شکل متفاوتی از نویز را ایجاد کند.

پرتوهای کیهانی، سیگنال‌های مخابراتی، تابش‌های رادیواکتیو و بسیاری دیگر نیز توانایی ایجاد نویز در دوربین‌های دیجیتال و آنالوگ را دارند اما معمولاً بسیار ناچیز هستند. [۱۰]

۲-۳-۱- مقدمه‌ای بر تاری

تاری یک نوع خرابی در تصویر می‌باشد که منجر به سرریز شدن اطلاعات یک پیکسل به پیکسل‌های مجاور می‌شود. تاری در تصویر می‌تواند در چندین جهت اتفاق بیفتد؛ مثلاً چنانچه تاری تصویر ناشی از عدم تنظیم لنز دوربین و یا تمرکز بر روی بخشی از تصویر باشد، در این صورت سرریز اطلاعات در کلیه‌ی جهات در پیکسل‌های همسایه اتفاق می‌افتد، اما اگر تاری به دلیل حرکت دوربین اتفاق بیفتد در آن صورت سرریز اطلاعات فقط در راستای حرکت اتفاق می‌افتد. [۱۱] و [۱۲] به‌طور کلی می‌توان گفت که با تار شدن یک

تصویر، اختلاف میان شدت روشنایی پیکسل‌های مجاور کم می‌شود، در نتیجه در اثر تاری اطلاعات فرکانس بالای تصویر کاهش یافته و دیدن جزئیات در تصویر دچار مشکل می‌شود. [۱۳]

تاری در تصویر به دلایل زیادی از جمله عدم تنظیم فاصله کانونی، لرزش دوربین و حرکت ایجاد می‌شود. بیشتر تحقیقات صورت گرفته در زمینه تاری بر روی شناسایی و رفع تاری صورت گرفته است که در آن‌ها میزان تاری به همه پیکسل‌های تصویر یکسان فرض می‌شود. [۱۴]

میزان تاری در همه‌ی تصویر یکسان نیست یعنی متغیر با مکان می‌باشد، در نتیجه برای رفع تاری تصویر بایستی ابتدا از میزان تاری پیکسل‌های آن اطلاع پیدا کرد. [۱۵-۱۸] می‌توان گفت که یکی از مسائل چالش برانگیز برای محققان در زمینه‌ی پردازش تصویر، ارزیابی کیفیت تصویر می‌باشد، هدفی که محققان از ارزیابی کیفیت تصویر دنبال می‌کنند، طراحی و توسعه‌ی الگوریتم‌ها و همچنین سنجشی برای تشخیص میزان تاری و وضوح (تیزی) در تصویر می‌باشد.

همان‌طور که پیش‌تر اشاره کردیم تاری در تصویر به دلایل زیادی از جمله عدم تنظیم فاصله کانونی، لرزش دوربین، حرکت ایجاد می‌شود. یکی از کارهای چالش برانگیز طراحی یک الگوریتم به منظور محاسبه‌ی اندازه‌ی کیفیت تصویر به صورت کورکورانه می‌باشد. در برخی از برنامه‌های پردازش تصویر بسیار مهم است که بتوان میزان تاری تصاویر را تعیین کیفیت کرد. [۱۹]

از مشکلات عمده درگیر در امر تصویربرداری، مسئله عدم وضوح تصاویر می‌باشد. عدم وضوح تصاویر یا اصطلاحاً (محوشدگی) در تصاویر هنگامی اتفاق می‌افتد که به دلایلی ناخواسته تشخیص سوژه‌ای که ما از آن تصویر گرفته‌ایم، مشکل باشد. [۲۰] لازم به ذکر است که در اینجا منظور از محوشدگی همان تاری در تصویر است. این محوشدگی یا تاری در تصاویر به دو صورت نمود پیدا می‌کند؛ یا کل تصویر محوشده است، یعنی در هیچ کجای تصویر هیچ قسمتی دارای وضوح کامل نیست که این مشکل نشان‌دهنده لرزش دوربین در هنگام عکس‌برداری است و یا بخشی از تصویر واضح است، اما این بخش سوژه اصلی ما نیست. در این حالت، احتمالاً یا سوژه متحرک بوده یا اینکه فوکوس دوربین درست انجام نشده است. انواع مختلف تاری تصویر یا به عبارتی علل مختلف تاری [۲۱] عبارت است از:

• لرزش دوربین

- حرکت صحنه

- عدم فوکوس صحیح

(الف)



(ب)



(ج)



شکل ۲-۳ انواع مختلف تاری در تصویر، (الف) نشان‌دهنده تاری ناشی از لرزش دوربین، (ب) نشان‌دهنده تاری ناشی از

سوژه متحرک، (ج) نشان‌دهنده تاری ناشی از فوکوس نادرست (فرض بر این است که سوژه توپ عقبی است)

۲-۳-۲- لرزش دوربین در چه مواقعی موجب تاری می‌شود؟

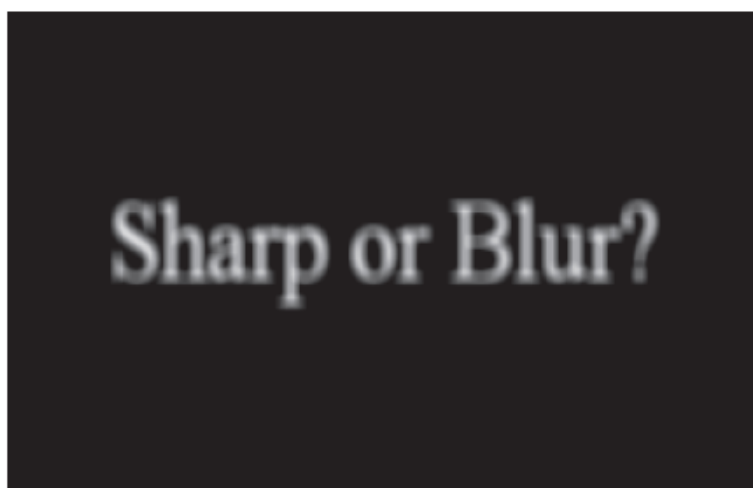
- در مکان‌هایی که نور کم است: در چنین حالتی برای گرفتن تصویر، نیاز است تا سرعت شاتر دوربین

را کم کرد تا زمان نوردهی تصویر افزایش پیدا کند، در حدود چند ثانیه باشد (تا نور کافی برای

ثبت تصویر فراهم شود). در این حالت یک لرزش کوچک در دوربین باعث تار شدن تصویر می‌شود.

- استفاده از لنزهای تله فوتو: در این حالت نیز هنگامی که روی یک صحنه‌ای زوم شده است. یک لرزش کوچک در دوربین باعث ایجاد تاری در تصویر می‌شود زیرا کوچک‌ترین لرزشی در دوربین باعث جابجایی زیادی در سطح تصویر می‌شود.

اولین مشاهده‌ی ما از تصویر بستگی به محتوای مکانی تصویر دارد اگر به یک تصویر که شامل یک قسمت تار شده پیرامون یک ناحیه‌ی همگن می‌باشد نگاه کنیم، حتی اگر بخش کوچکی از این عکس تار باشد، فکر می‌کنیم که آن یک تصویر غیرواضح است. به این دلیل در تجزیه و تحلیل تغییرپذیری پیکسل‌های مجاور، ما فقط به حساب پیکسل‌هایی می‌پردازیم که بعد از مرحله‌ی تاری تغییر کرده‌اند. [۲۲]



شکل ۲-۴ یک تصویر ساده که در سطحی بزرگ تار و بلوره شده باشد

علاوه بر این، با تار کردن یک لبه‌ی تیز روی یک ناحیه یک‌دست، مشاهده می‌کنیم که تغییر اندکی روی پیکسل‌های مجاور پیرامون لبه به وجود آورده‌ایم که روی تصویر واضح وجود ندارد برای جلوگیری از بروز این تغییرات فقط تغییرات پیکسل‌های مجاور را که بعد از مرحله تاری کاهش داده شده‌اند در نظر می‌گیریم. در حقیقت زیاد یا کم شدن نویزهای افزایشی با تصویر تار با وضوح بالا یا کم ظاهر می‌شوند. با در نظر گرفتن تمام تغییرات تصویر که بعد از مرحله‌ی تاری کاهش یافته‌اند. همچنین پیکسل‌هایی که شامل اطلاعات نویزی می‌شوند را تحلیل کردیم که می‌تواند بر روی لبه‌ها در مناطق منسجم و یا مناطق همگن قرار گیرد. روی یک تصویری که از یک بعد تار شده است، چشم تأثیر تاری را فقط در آن بعد درک می‌کند

اما تیزی در جهات دیگر را درک نمی‌کند. [۲۳ و ۲۴] به این دلیل، تاری در جهات افقی و عمودی را تخمین می‌زنیم و تاری را که بیشتر آزاردهنده باشد، به عنوان تاری نهایی انتخاب می‌کنیم.

۲-۳-۳- تاری

اگر که بتوانیم در یک تصویر تمام اشیا و اشکال را با وضوح بالا ببینیم آن تصویر از وضوح بالایی برخوردار است. به عنوان مثال تصویری که چهره در آن قرار دارد اگر که قادر به تشخیص چشم‌ها، گوش‌ها، پیشانی، بینی، لب‌ها باشیم، تصویر وضوح بالایی دارد. این شکل از اشیا از لبه‌های آن ناشی می‌شود؛ بنابراین در تاری مقدار لبه‌ها را تقلیل می‌کنیم و تغییری اندک از یک رنگ به دیگر رنگ‌ها ایجاد می‌کنیم.

۲-۳-۳-۱- تاری در برابر زوم کردن

زمانی که یک تصویر با استفاده از تکرار پیکسل‌ها زوم شود و عوامل زوم افزایش یابد تصویر تار می‌شود. [۲۵] این تصویر جزئیات کمتری دارد اما تاری آن درست نیست، زیرا در تاری پیکسل‌های جدیدی به یک تصویر اضافه می‌شود که باعث افزایش تعداد کلی پیکسل‌ها در یک تصویر می‌گردد، در صورتی که در زوم کردن تعداد پیکسل‌های یک تصویر نرمال و تصویر تار شده یکسان باقی می‌ماند. [۲۶]

۲-۳-۳-۲- انواع فیلترها

تاری توسط روش‌های متفاوتی می‌تواند به دست آید. یک روش معمول آن استفاده از فیلترهایی می‌باشد که منجر به تاری می‌گردد.

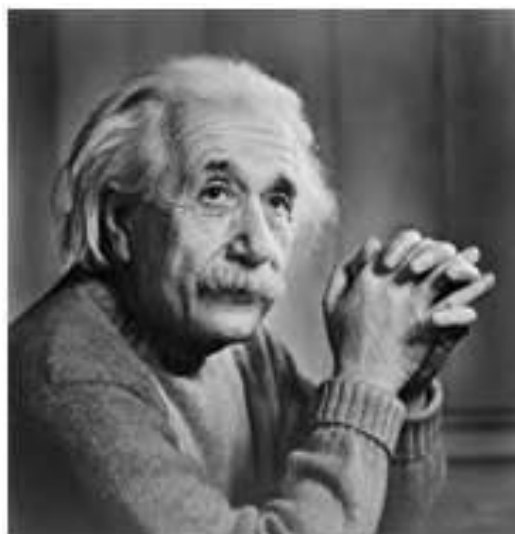
- فیلتر میانگین^۱
- فیلتر میانگین وزنی^۲
- فیلتر گاوسی^۳

^۱ Mean filter

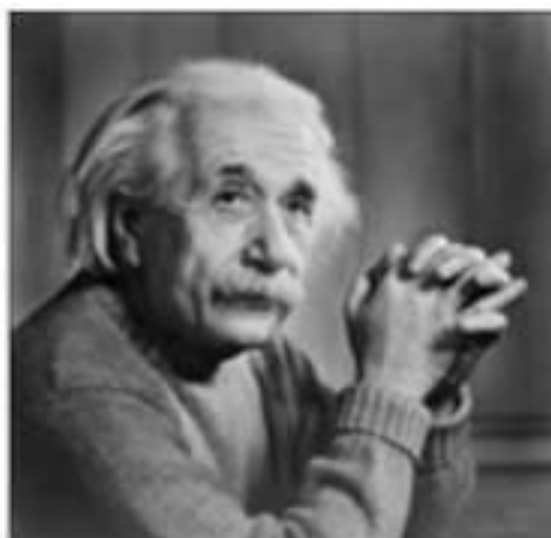
^۲ Weighted average filter

^۳ Gaussian filter

(الف)



(ب)



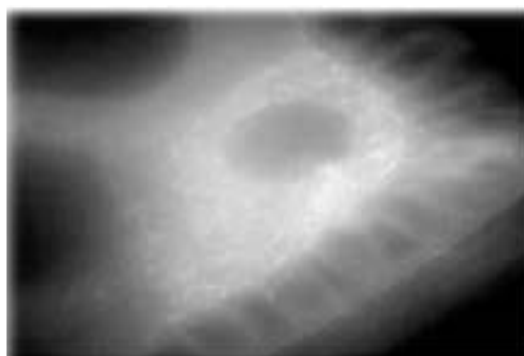
شکل ۵-۲ مقایسه میان تصویر اصلی (الف) و تصویر تار شده (ب) پس از اعمال ماسک

ممکن است که نتایج به دست آمده واضح به نظر نرسد، در این صورت میزان تاری را افزایش می دهیم. تاری می تواند با افزایش سایز ماسک ها به دست آید. هر چه سایز ماسک بزرگ تر باشد، میزان تاری افزایش می یابد؛ زیرا با ماسک بزرگ تر تعداد بیشتری از پیکسل ها فراهم می آیند و تغییر حالت نرمی صورت می گیرد.

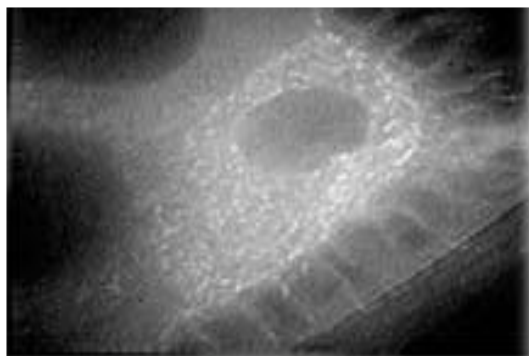
۲-۳-۳- کاربردهای حذف تاری

به طور مثال اگر از خودرویی عکس گرفته شده باشد و در آن تصویر پلاک به قدری بوده که قابل تشخیص نباشد و یا تصاویر نجومی و پزشکی که به دلیل استفاده از میکروسکوپ های الکترونی و تصویربرداری از ذرات بسیار ریز، امکان جلوگیری از پدیده ی تاری وجود ندارد. همچنین کاربرد دیگری که بسیار مورد مشاهده نیز می باشد دوربین های دیجیتال شخصی می باشد، در حقیقت بیشترین علتی که موجب تاری می شود در مواقعی است که به هر علتی نمی توان از لرزش دوربین در محیط جلوگیری کرد.

(الف)



(ب)



(ج)



(د)



شکل ۲-۶ حذف تاری در تصاویر میکروسکوپی نجومی، (الف و ج) تصاویر تار گرفته شده توسط میکروسکوپ و تلسکوپ.

(ب و د) همان تصاویر بعد از حذف تاری

۲-۴- الگوریتم‌های کیفیت تصویر

روش‌های اندازه‌گیری میزان تنزل کیفیت تصویر به دو روش صورت می‌گیرد: [۲۶-۲۹]

- به صورت ذهنی و با در نظر گرفتن امتیازات بصری انسان

- به صورت آماری یا عددی

روش‌های عددی مبتنی بر اساس اندازه‌گیری‌های ریاضی‌اند، درحالی‌که روش‌های ذهنی به درک گروهی از افراد بیننده از جمله حرفه‌ای‌ها و یا غیر حرفه‌ای‌ها تکیه می‌کنند. ارزیابی آماری به عنوان ابزاری قابل اطمینان برای دسته‌بندی کیفیت تصویر کمتر ترجیح داده می‌شود، از طریق ارزیابی بصری کیفیت تصویر توسط انسان تخمین زده می‌شود به هر حال اداره‌ی یک ارزیابی بصری کار آسانی نیست زیرا که هزینه‌بر و زمان‌گیر است. [۳۰] با توجه به محدودیت در ارزیابی بصری محققان تلاش می‌کنند تا روش‌های آماری طراحی

کنند که محصولاتی نزدیک به بینایی انسان داشته باشد. در بسیاری از کارها در حوزه پردازش تصویر به سبب محدودیت‌های موجود از روش‌های عینی برای ارزیابی کیفیت تصاویر و ویدیو استفاده می‌شود. همانند روش‌های مورد نظر در این تحقیق که برای درک میزان تثبیت و پایداری تصاویر ویدیو نیاز به درک و اعمال نظر بیننده و قضاوت عینی انسان، چه خبره یا غیرخبره، می‌باشد.

۲-۴-۱- روش‌های عینی ارزیابی کیفیت تصاویر

- ارزیابی کیفیت تصویر تمام مرجع^۱
- ارزیابی کیفیت تصویر نیمه مرجع^۲
- ارزیابی کیفیت تصویر فارغ از مرجع^۳

یک ارزیابی تمام مرجع نیازمند در دسترس بودن تصاویر اصلی می‌باشد درحالی‌که روش‌های فارغ از مرجع کیفیت بصری تصاویری را ارزیابی می‌کنند که متأثر از اغتشاشات بخصوصی هستند و روش‌های نیمه مرجع نیز به بخشی از اطلاعات تصویر مرجع نیاز دارند و در حقیقت یک برآیند میان دو روش تمام مرجع و فارغ از مرجع می‌باشند. [۲۸و۳۰]

با افزایش تقاضا در برنامه‌های مبتنی بر تصویر، ارزیابی کیفیت تصویر قابل اعتماد و کارآمد، اهمیت یافته است. اندازه‌گیری کیفیت تصویر اهمیتی اساسی برای بسیاری از برنامه‌های پردازش تصویر دارد. هدف روش‌های ارزیابی کیفیت تصویر، ارزیابی اتوماتیک کیفیت تصویر در راستای نظرات و قضاوت انسان می‌باشد. [۳۱و۳۲] در ۸-۲ در این باره به تفصیل توضیح داده خواهد شد.

۲-۵- برخی از روش‌های حذف نویز

۲-۵-۱- فیلتر وینر

فیلتر وینر معمولاً برای بازیابی خطی تصاویر خراب شده مورد استفاده قرار گرفته و بر مبنای حداقل مربعات خطا (MSE) عمل می‌کند. درعین حال، این فیلتر یک فیلتر خطی بهینه برای محیط ایستا است که بهترین

¹ Full-Reference Image Quality Assessment (FR)

² Reduced-Reference Image Quality Assessment (RR)

³ No-Reference Image Quality Assessment (NR)

تخمین از سیگنال را تولید می‌کند. در اینجا، کلمه بهینه در خصوص حداقل میانگین مربعات خطا استفاده می‌شود. فرضیه اساسی در نظریه فیلتر وینر این است که طیف توان تصویر ایده‌آل و نویز معلوم هستند. [۳۳]

نشان داده شده است که فیلتر وینر امکان تخمین تصویر اصلی را به عنوان جواب مسئله خطی دارد، البته زمانی که تصاویر اندازه‌گیری شده دست‌کاری شود. این تخمین را می‌توان به صورت کانولوشن سیگنال اندازه‌گیری شده با یک فیلتر ثابت (نامعتبر) بیان نمود. پاسخ فرکانس فیلتر وینر غیر علی (زمانی که به مقدار زیادی از داده‌ها نیاز داریم بنابراین این فیلتر برای کاربردهای real-time مناسب نیست) برابر است با فرمول (۱-۲)

$$H(w) = \frac{P_s(w)}{P_s(w) + P_N(w)}$$

در این رابطه $P_s(W)$ و $P_n(W)$ به ترتیب بیانگر چگالی طیف توان سیگنال صحیح و نویز هستند. [۳۴]

۲-۵-۲- فیلتر میانه

به دلایل زیر فیلترهای خطی جذاب هستند: آن‌ها ویژگی‌های جبری مفیدی را مطرح می‌کنند، عملیات بکار رفته در آن‌ها به سادگی قابل فهم است، از طریق تبدیل فوریه، رابطه مستقیمی با پاسخ فرکانس و ویژگی‌های آماری آن‌ها دارند.

فیلتر کردن میانه توسط Tukey برای یکنواخت کردن سری‌های زمانی اقتصادی معرفی شد. فیلتر یک‌بعدی را می‌توان با لغزاندن یک پنجره (اطراف پیکسل‌های مورد نظر) بر روی تمام تصویر پیاده‌سازی نمود. نتیجه عملکرد میانه بر روی پیکسل‌های درون پنجره همان خروجی فیلتر است. فرض کنید که N نشان‌دهنده سائز پنجره باشد و چون سائز پنجره باید یک عدد صحیح فرد باشد، پس $N=2k+1$. عبارت ریاضی فیلتر میانه با طول N را می‌توان به صورت فرمول (۲-۲) بیان نمود:

$$y(n) = MED[x(n-k), \dots, x(n), \dots, x(n+k)]$$

$X(n)$ و $y(n)$ به ترتیب ورودی و خروجی هستند. این اغلب میانه جاری یا مداوم نامیده می‌شود. در پردازش تصویر، معمولاً فیلتر میانه با حرکت یک پنجره مربعی یا متقاطع شکل و انتخاب میانه مقادیر پیکسل درون خروجی پنجره جاری اعمال می‌شود.

۲-۵-۳- فیلتر گوسی

در مسائلی که با ساختارهای بلورینگ سروکار داریم، از فیلتر گوسی برای حذف سایه و هموارسازی تصویر استفاده می‌کنیم. در یک فضای دوبعدی، کرنل گوسی را می‌توان به صورت فرمول (۲-۳) تعریف نمود:

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

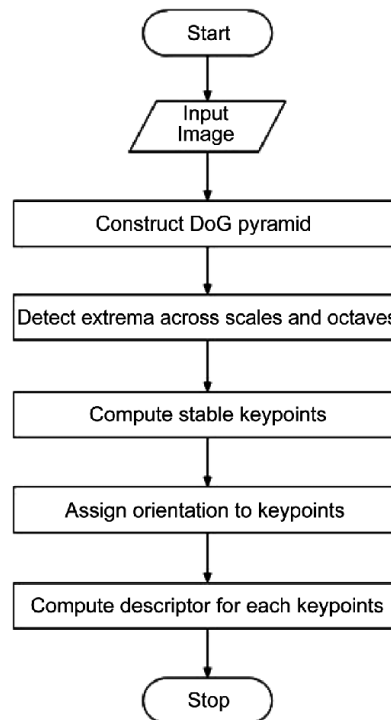
فیلتر گوسی راه‌حل بهتری برای فیلتر کردن پایین گذر تصویر است. از جمله مزایای این فیلتر می‌توان گفت این فیلتر یک فیلتر میانگین‌گیر موثر است و متوسط وزن دار مقادیر درون پنجره محلی را محاسبه می‌کند. [۳۵ و ۳۶]

۲-۵-۴- الگوریتم تبدیل ویژگی مقیاس ثابت^۱

الگوریتم تبدیل ویژگی مقیاس ثابت یکی از متداول‌ترین روش‌های مبتنی بر ویژگی است که در برابر تغییرات مقیاس و چرخش ثابت است و در برابر تغییرات روشنایی و تبدیلات افین^۲ و نویز نیز پایدار می‌ماند. این مزایا سبب اهمیت این الگوریتم در فرایند تناظریابی شده است اما ماهیت پیچیده تصاویر سنجش از دوری سبب ایجاد تناظرهای نادرست بسیار زیادی می‌شود. تحقیقات بسیاری به منظور بهبود دقت تناظریابی در مرحله استخراج ویژگی و توصیفگرهای آن انجام شده است. مرحله استخراج ویژگی در این الگوریتم خود شامل سه مرحله است: استخراج اکستریم‌های فضای مقیاس، بهبود دقت موقعیت و حذف اکستریم‌های ناپایدار و در آخر، تخصیص جهت به هر ویژگی ایجاد شده. در ادامه جزئیات هریک از این مراحل را شرح می‌دهیم. [۳۷ و ۳۸]

^۱ SIFT (Scale-invariant feature transform)

^۲ Affine



شکل ۲-۷ فلوچارت SIFT

۲-۵-۵- کاهش نویز در تصاویر با استفاده از روش میانگین‌گیری

در این روش فرض بر این است که محل قرارگیری نویزها در تصویر کاملاً به صورت تصادفی است. لذا نوسانات تصادفی که در بالا و پایین اطلاعات اصلی تصویر ایجاد شده و باعث ایجاد نویز شده‌اند با میانگین‌گیری و ترکیب چند عکس رو به کاهش نهاده و به تدریج هر چه تعداد عکس بیشتری را ترکیب کنیم به سطح واقعی خود نزدیک‌تر شده و یکسان می‌شوند. [۳۹]

۲-۶- سوابق کار

برای رفع مشکل لرزش ناخواسته و جابجایی فریم‌ها روش‌های متعددی معرفی شده است. در حالت کلی روش‌های مبتنی بر تکنیک‌های پردازش تصویر عموماً شامل سه مرحله هستند [۱۱ و ۴۰]:

تخمین حرکت^۱

جبران حرکت^۲

^۱ Motion estimation

^۲ Motion compensation

ترکیب تصویر^۱

هدف الگوریتم های پردازش ویدئو که در جهت پایداری سازی تصویر توسعه یافته اند، حذف حرکت های ناخواسته دوربین است.

سرعت مناسب و درعین حال دقت بالای این الگوریتم ها در مواجهه با شرایط نویزی و وجود لرزش، از اهمیت بالایی برخوردارند. تفکیک صحیح حرکت های ناخواسته از حرکت های عمدی و همچنین نرمی تصویر خروجی نیز از جمله اهدافی است که در نهایت می بایست به بهترین شکل انجام شود. ساز و کارهای سخت افزاری متعددی چه به صورت مکانیکی و چه به صورت الکتریکی برای پایداری سازی ویدئو توسعه یافته اند. [۴۱]

اما تجهیزات مکانیکی عموماً دقیق نیستند و از طرفی بسیار حجیم و سنگین هستند. تجهیزات الکتریکی نیز از نظر انعطاف پذیری، بسیار محدودیت دارند. راهکارهای دیجیتال می توانند به خوبی این مشکلات را برطرف سازند [۴۴ و ۴۵]؛ اما مشکل اصلی این الگوریتم ها پیچیدگی بالای زمانی آن ها و ضعف در تخمین حرکت عمدی دوربین می باشد. یک راهکار دیجیتال جدید ارائه شده [۳۵ و ۴۶] است که با بهره گیری از فیلتر ذره، حرکت عمدی دوربین را به صورت برخط و در عین حال به صورت هدفمندتری پیش بینی می کند. همچنین با تعیین تطبیقی پارامترهای فیلتر ذره، تعداد فیلترهای ذره و واریانس نویز سیستم، سرعت الگوریتم پیشنهادی بهتر شده و عملیات تخمین حرکت عمدی دوربین نیز به صورت مناسب تری انجام می گیرد. [۴۷-۵۰] از فیلتر کالمن جهت تعیین پارامترهای حرکت دوربین بهره گرفته اند؛ اما فیلتر کالمن در مواجهه با مدل های غیرخطی با نویز غیر گاوسی به خوبی عمل نمی کند. [۵۱-۵۳]

از فیلتر ذره برای تشخیص پارامترهای حرکت عمومی دوربین استفاده کرده اند ولی همگی جهت تشخیص حرکت عمدی دوربین از فیلتر کالمن بهره گرفته اند. برخی روش ها [۵۴] از فیلتر پایین گذر علی استفاده کرده اند که در پیش بینی حرکات عمدی دوربین در تغییرات ناگهانی، چندان مناسب عمل نمی کند. همچنین برخی دیگر [۵۵ و ۵۶] از منطق فازی جهت پایداری سازی ویدئو استفاده و نشان داده شده که منطق فازی سرعت و کارایی بالایی در کاربرد پایداری سازی ویدئو از خود نشان می دهد.

¹ Image composition

مدل‌های حرکتی بسیاری برای حرکت دوربین پیشنهاد شده‌اند. مدل افاین محدود شده با چهار پارامتر [۵۷]، علیرغم سادگی، با ترکیب با فیلتر ذره در کاربرد پایدارسازی ویدئو، کارایی و مناسبی دارد. بر

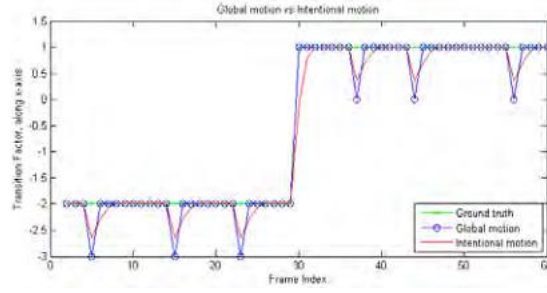
اساس این مدل حرکتی دو بعدی، انتقال نقطه (x, y) به نقطه (y', x') به شکل زیر قابل بیان است:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \cos \varphi & -\sin \varphi & T_x \\ \sin \varphi & \alpha \cos \varphi & T_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \text{فرمول (۴-۲)}$$

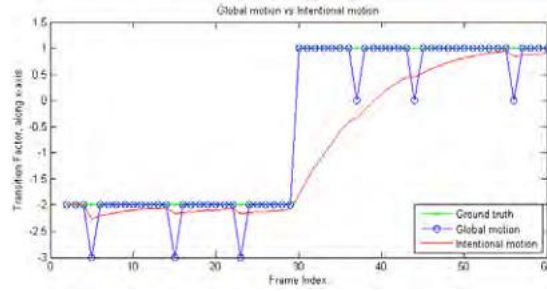
که T, φ, α و T به ترتیب پارامترهای بزرگنمایی، دوران و جابجایی در راستای محورهای x و y هستند. با در نظر گرفتن این مدل حرکتی و در اختیار داشتن پارامترهای حرکت عمومی دوربین، حال می‌بایست حرکت‌های ناخواسته از حرکت‌های عمدی دوربین تمیز داده شوند. از آنجایی که این امر می‌بایست به صورت برخط انجام شود، می‌توان از یک فیلتر پایین گذر علی برای نرم کردن حرکت عمومی و تخمین حرکت عمدی دوربین بهره گرفت. مبنای کار این فیلتر که در واقع یک میانگین‌گیری وزن‌دار در یک بازه زمانی مشخص T ، روی پارامترهای حرکت عمومی دوربین می‌باشد، بر این اساس استوار است که حرکت‌های ناخواسته عموماً فرکانس بالاتری دارند و نمی‌بایست از فیلتر عبور کنند. رابطه زیر، این فیلتر را نشان می‌دهد:

$$\theta_t^d = \sum_{i=1}^T w_i * \theta_{t-i}^g \quad \text{فرمول (۵-۲)}$$

اما مشکل این روش این است که چنانچه پنجره T ، کوچک در نظر گرفته شود، قدرت نرم کردن این فیلتر و یا به عبارت دیگر قدرت حذف حرکت‌های ناخواسته دوربین، مطابق آنچه در شکل (۱-الف) قابل مشاهده می‌باشد، به شدت کاهش می‌یابد. در مقابل با بیشتر کردن این پارامتر مطابق شکل (۱-ب)، تخمین حرکت عمدی دوربین در تغییرات ناگهانی، دچار مشکل می‌شود. در ادامه سعی می‌شود، با تغییر مفهوم حذف حرکت‌های ناخواسته از نرم کردن حرکت عمومی به پیش بینی حرکت عمدی دوربین به کمک فیلتر ذره فازی، پارامترهای حرکت عمدی دوربین به شکل مناسب‌تری برآورد گردد. [۵۸]



(الف)



(ب)

شکل ۲-۸ تخمین T_x از مدل حرکت عمدی دوربین با استفاده از فیلتر پایین گذر علی با $T=3$ (الف) و $T=30$ (ب) [۳۵]

با در نظر گرفتن مدل حرکت عمدی دوربین به عنوان یک سیستم پویا و پارامترهای مدل حرکت عمومی دوربین برای فریم t ، به عنوان مشاهدات حاوی نویز این سیستم، مسأله مورد نظر می‌تواند در قالب مسأله پیش‌بینی وضعیت ناشناخته سیستم در لحظه t یعنی $\theta_t = [\alpha, \phi, T_x, T_y]$ ، در نظر گرفته شود. همچنین فرض می‌شود، مدل تغییر وضعیت، E_t و مدل مشاهدات O_t ، به شکل زیر تعریف می‌شوند:

$$\begin{aligned}\theta_t &= E_t(\theta_{t-1}, U_{t-1}) \\ Y_t &= O_t(\theta_t, N_t)\end{aligned}$$

در رابطه فوق U_t نویز سیستم و N_t نویز مشاهدات می‌باشند و به صورت توابع گاوسی در نظر گرفته شده‌اند. یک ذره یک نمونه وزن‌دار است که می‌تواند چگالی پیشین را تخمین بزند [۲۰-۲۳]. هر ذره $P_t = \{\hat{\theta}_t, w_t\}$ با وضعیتی که تخمین می‌زند و وزن متناظر با آن مشخص می‌شود. وزن w_t ، متناسب با احتمال پسین $P(Y_t | \hat{\theta}_t)$ نرمال می‌باشد. $Y_t = [v_t^1 \dots v_t^n]$ بردار مشاهدات در لحظه t و در مسأله مورد نظر شامل پارامترهای مدل حرکت عمومی نویزی دوربین می‌باشد. $Y_{1:t}$ پارامترهای مدل حرکت عمومی دوربین از لحظه اول تا لحظه t می‌باشد.

به منظور مقداردهی اولیه فیلترهای ذره، در ابتدا K نمونه $\{\theta_t^i\}_{i=1}^K$ از تابع چگالی $P(\theta_{t-1}|Y_{1:t})$ نمونه از تابع $\{U_{t=1}^i\}$ احتمال نویز U_{t-1} و همچنین K نمونه نویز $\{N_t^i\}_{i=1}^K$ از تابع احتمال N_t نمونه برداری می‌شود. با اعمال این نمونه‌ها در روابط وضعیت θ_t در اختیار خواهند بود. سپس وزن هر ذره به کمک این مقادیر طبق رابطه زیر محاسبه می‌گردد:

$$w_t^i \propto \frac{P(Y_t^i|\theta_t)P(\theta_t|\theta_{t-1})}{g(\theta_t|\theta_{t-1}, Y_{1:t})} \quad \text{فرمول (۶-۲)}$$

که در رابطه فوق g تابع توزیع پیشنهادی می‌باشد. در تحقیق دیگری [۳۴] از روشی مشابه با الگوریتم نمونه‌برداری مجدد ترتیبی [۱۸ و ۱۹]، استفاده شده است که با تقریب $P(\theta_t|\theta_{t-1}) \approx g(\theta_t|\theta_{t-1}, Y_{1:t})$ ، رابطه قبل را به شکل زیر ساده می‌سازد:

$$w_t^i \propto P(Y_t^i|\theta_t^i)$$

سپس وزن هر ذره با رابطه زیر نرمال‌سازی می‌شود:

$$w_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i} \quad \text{فرمول (۷-۲)}$$

در نهایت، وضعیت تخمین زده شده توسط جمعیت ذرات در زمان t ، به شکل زیر تعیین می‌گردد:

$$\hat{\theta}_t = \operatorname{argmax}_{\theta_t} P(\theta_t|Y_{1:t}) \approx \operatorname{argmax}_{\theta_t} w_t^i \quad \text{فرمول (۸-۲)}$$

جهت تشکیل مناسب‌تر جمعیت ذرات، نقش دو پارامتر تعداد ذرات و واریانس نویز سیستم در هر مرحله، بسیار مهم می‌باشد. هرچه تغییرات پارامترهای حرکت عمومی دوربین بیشتر بوده و این تغییرات تداوم داشته باشد، به تعداد ذرات بیشتری برای پوشش فضای جستجو نیاز خواهد بود. در غیر این صورت با کاهش جمعیت ذرات میتوان سرعت الگوریتم پیشنهادی را افزایش داد. از طرف دیگر، چنانچه این تغییرات زیاد بوده و تداوم یابد، لازم خواهد بود تا با افزایش نویز سیستم، توزیع مناسب‌تری از ذرات در فضای جستجو انجام گیرد.

در غیر این صورت با کاهش این نویز میتوان دقت پیش‌بینی‌ها را افزایش داد.

بنابراین با ارائه یک راهکار فازی ساده، با حداقل قوانین مورد نیاز (۹ قانون) که در جدول ۱-۲ آورده شده است، تعداد ذرات و واریانس نویز سیستم به صورت تطبیقی، محاسبه می‌شوند. این سیستم کنترل فازی دو ورودی و یک خروجی دارد که ورودی اول، اختلاف میان پارامتر پیش‌بینی شده حرکت عمدی با پارامتر حرکت عمومی دوربین و ورودی دوم تداوم این تغییر را ارزیابی می‌کند و به شکل زیر قابل تعریف هستند:

$$\begin{aligned} input1_t &= |\hat{\theta}_t - \theta_t| \\ input2_t &= |input1_t - input1_{t-1}| \end{aligned}$$

تعیین پارامترهای فیلتر ذره با کمک این سیستم کنترل فازی موجب می‌شود، حرکت های ناخواسته دوربین به شکل بهتری حذف شده و در عین حال حرکت های عمدی دوربین به شکل اصلی خود حفظ شوند و تحت عنوان حرکت ناخواسته دوربین، حذف نگردند.

جدول ۱-۲ قوانین سیستم کنترل فازی

		Input 1			
		Z	P	B	
Input 2	Z	Z	Z	P	Z = Zero P = Positive B = Big-Positive
	P	P	P	P	
	B	P	P	B	

جهت وزن دهی ذرات نیز از فاصله اقلیدسی استفاده می‌شود.

هرچه تخمین ذره مورد نظر به وضعیت واقعی نزدیک تر باشد، وزن بیشتری خواهد داشت. وزن هر ذره به شکل زیر قابل محاسبه می‌باشد:

$$w_t^i \propto \frac{1}{||\hat{\theta}_t - \theta_t||}$$

پس از محاسبه مدل حرکت بین فریمی عمومی و عمدی دوربین، با درنظرگرفتن یک فریم، مثلاً فریم اول، به عنوان فریم مرجع، جبران حرکت به شکل زیر قابل انجام است:

$$X_{t,i}^s = (A_{t-1}^d * \dots * A_1^d) * (A_{t-1}^g * \dots * A_1^g)^{-1} * X_{t,i} \quad \text{فرمول (۹-۲)}$$

$X_{t,i}^s$ و $X_{t,i}$ به ترتیب مختصات پیکسل i از فریم ناپایدار و پایدار t و Ag و Ad به ترتیب ماتریس

افاین حرکت عمومی و عمدی دوربین می‌باشند. [۳۵]

در تحقیق انجام شده توسط Shen و همکاران [۹] ابتدا مکان نقاط کلیدی با استفاده از بلاک مدور یافته شده سپس تبدیل نسبت سلبی^۱ توسط یک روش مبتنی بر تخمین چندجمله‌ای هموار می‌شود؛ اما این روش حتی در دیتاست تصاویر 300×216 پیکسل، تنها به سرعت ۱۰fps (فریم بر ثانیه) دست یافت. Lee و همکارانش [۱] در سال ۲۰۰۹ روشی ارائه کردند که در آن با توجه به مسیرهایی^۲ که از ویژگی‌های داخل چند فریم متوالی استخراج می‌شد، استفاده می‌کرد. این روش برای کم کردن میزان فضای خالی فریم‌ها پس از پایدارسازی، میزان پایدارسازی را کمتر کرده‌اند تا حدی که از لحاظ بصری قابل قبول به نظر آید؛ اما در مواردی مانند تصاویر با تغییر عمق زیاد، عملکرد روش کم می‌شد.

در سال ۲۰۱۳ نیز روشی [۴] ارائه شد که با استفاده از الگوریتم SIFT به یافتن نقاط کلیدی برای هر دو فریم متوالی و فیلترهای کالمن و Median برای رفع نویز تصاویر کمک گرفتند. این پیاده‌سازی به‌طور خاص بر روی تصاویر هوایی اعمال و برای ارزیابی از معیار PSNR^۳ و ITF^۴ که مبتنی بر معیار ذکر شده اخیر است و مقادیر این دو معیار هرچه بیشتر باشند نشانگر میزان پایداری بیشتر تصاویر است، استفاده کردند. همچنین پیشنهاد استفاده از Optical flow برای تخمین جابجایی‌های محلی را برای تحقیقات آینده ارائه کردند.

در روشی [۳۹] که در سال ۲۰۱۶ انجام شد با استفاده از یک فیلتر ترکیبی به پایدارسازی تصاویر پرداختند. در ابتدا نقاط حاشیه‌ای^۵ با روش GFTT^۶ استخراج و از آن‌ها برای محاسبه Optical flow برای دو فریم متوالی استفاده شد. در ادامه پارامترهای جابجایی رخ داده، تخمین زده می‌شوند که برای هموار کردن این پارامترها از ترکیب فیلترهای ذکر شده کمک گرفته می‌شود؛ یک فیلتر کالمن، یک فیلتر پائین‌گذر و استفاده مجدد از فیلتر کالمن برای حذف تمام حرکت‌های ناخواسته. فرضی برای روش‌های مبتنی بر فیلتر کالمن وجود دارد [۴۰] و آن دخالت نویز گوسی سفید^۷ به‌تنهایی در درون تصاویر است که با این روش این فرض

¹ Affine transformation

² Trajectories

³ Peak Signal-to-Noise Ratio

⁴ Interframe Transformation Fidelity

⁵ Corner points

⁶ Good feature to track

⁷ White Gaussian noise

از بین می‌رود. معیار MSE^1 که برای دو فریم متوالی محاسبه شده، بیانگر میزان بهبود حدوداً ۱۰ درصدی نسبت به حالت استفاده از فیلتر کالمن به‌تنهایی می‌باشد.

روشی دیگر [۴۱] در سال ۲۰۱۶ معرفی شده است که استفاده از $SURF^2$ را به دلیل پیچیدگی کمتر محاسباتی جایگزین SIFT کرده و همچنین ادعا شده است که روشی باثبات‌تر و تکرارپذیرتر از روش‌های مبتنی بر یافتن نقاط حاشیه‌ای می‌باشد. برای هموارسازی پارامترهای حرکتی نیز از فیلتر پائین‌گذر استفاده کردند.

در روش [۳۴] که مبتنی بر رویکرد [۴۲] حذف اثر Phantom movement می‌باشد هدف افزایش مقاومت و سرعت و همچنین کاهش تغییر شکل فریم‌هاست. از تبدیلات هندسی^۳ برای تشریح ارتباط بین دو فریم متوالی و از SURF برای یافتن نقاط ویژگی بهره‌برده شده است. RANSAC نقاط جفت شده اشتباه را حذف کرده و با تبدیلات وابسته و بر پایه تبدیلات تصویری^۴ مسیرهای شکل‌گرفته بر اساس نقاط ویژگی را هموار کرده‌اند. برای تخمین جابجایی‌های عمودی از فیلتر باتروورث مرتبه دو^۵ استفاده شده که برای تمییز دادن با جابجایی‌های غیرعمودی حیاتی می‌باشد؛ اما این فیلترها در هنگام به دست آوردن خروجی دارای مقداری تأخیر می‌باشند و فیلتر مذکور نیز از این حیث مستثنا نیست. در مرحله بعد و با استفاده از یک مدل شبکه عصبی غیرخطی پارامترهای حرکتی خطی سازی شده و به فیلتر کالمن ارجاع داده می‌شوند تا میزان جابجایی‌های فریم تصحیح شده و فریم تغییر یافته شده حاصل شود. معیارهای سنچس ITF و RMSE معرفی شده و نتایج با سه الگوریتم مقایسه شده‌اند؛ روش [۴۳] مورد استفاده توسط وب‌سایت یوتیوب، روش [۴۲] قبلی ارائه شده توسط همین نویسندگان و نرم‌افزار تجاری Adobe after effects.

روش [۵] معرفی شده توسط Zerara از تصاویر ورودی زیر تصویری مربعی شکل و به مرکزیت محل چرخش تصویر به‌عنوان نمونه برداشته، آن‌ها را به بردارهایی 180×1 نگاشت می‌کند. مقادیر درایه‌ها میانگین مقادیر پیکسل‌ها در آن قطعه خطی تشکیل شده با سطر تصویر است. به ازای هر دو بردار برای فریم‌های متوالی،

¹ Mean Square Error

² Speeded-up robust features

³ Geometric transformation

⁴ Projective transformation

⁵ Second-order Butterworth

همبستگی متقابل یک‌بعدی^۱ را محاسبه کرده و با استفاده از آن زاویه چرخش را به دست آورده و فریم را پایدار می‌کند. این روش؛ اولاً فقط جابجایی‌های مربوط به چرخش فریم را تصحیح می‌کند دوماً نیاز به یافتن دقیق مرکز چرخش فریم‌ها دارد که علی‌رغم راه‌حل ارائه شده بازهم دارای خطاست.

همچنین در مطالعه‌ای [۷] روشی مطرح شده و با [۴۳] الگوریتم‌های مورد استفاده توسط وب‌سایت یوتیوب مقایسه شده است. این روش از ۷ مرحله تشکیل شده؛ با استفاده از SURF نقاط کلیدی استخراج و نقاط مرتبط باهم لینک می‌شوند، جابجایی‌های فریم تخمین زده شده و با الگوریتم RANSAC^۲ تشابهات غلط و حرکات مربوط به اجسام را حذف کرده و مسیرها شکل می‌گیرند، سپس این مسیرها با اعمال یک فیلتر پائین‌گذر (گوسی) هموار شده و میزان جابجایی شکل‌گرفته جبران می‌گردد، در انتها فریم برش می‌خورد تا تصاویر تثبیت شده حاصل شوند. معیار استفاده شده PSNR و ITF و دیتاست [۴۳] مورد استفاده نیز GaTech VideoStab انتخاب شده است. به‌طور متوسط نتایج فیلتر گوسی با فیلتر کالمن تفاوت زیادی نداشته اما بهبودی نسبت به تصاویر ورودی را نشان داد.

همچنین در [۶۷] توسعه یک الگوریتم برای مقابله با مشکل تثبیت‌کننده فیلم در زمان واقعی برای وسایل نقلیه هوایی بدون سرنشین (پهپاد) توضیح داده شده است. دو المان اصلی در الگوریتم وجود دارد: (۱) با طراحی یک مدل مناسب برای حرکت سراسری پهپاد، الگوریتم از لزوم تخمین عمومی‌ترین مدل حرکت جلوگیری می‌کند و مدل‌های حرکتی ساده‌تر را در نظر می‌گیرد. (۲) برای دستیابی به سرعت پردازش بالا، ردیابی مبتنی بر جریان نوری بجای روش‌های ردیابی معمولی و تطبیق استفاده شده توسط الگوریتم‌های پیشرفته هنر استفاده می‌شود.

در روالی مشابه ولی این بار بر روی زمین، [۶۸] تثبیت‌کننده فیلم‌های راهنمایی و رانندگی ضبط شده توسط دوربین‌های سوار شده وسیله نقلیه را در نظر می‌گیرد. در مقایسه با فیلم‌های ضبط شده توسط دوربین‌های دستی، فیلم‌های راهنمایی و رانندگی از چالش‌های بیشتری مانند صحنه‌های پویا، اشیاء پیش‌زمینه غالب و اختلاف‌منظر رنج می‌برند. از آنجایی که روش‌های معمولی تثبیت‌کننده فیلم معمولاً یک

¹ Cross-correlation

² Random sample consensus

هوموگرافی را فقط از مسیرهای ویژگی پس‌زمینه و قاب‌های تابشی با آن هم‌سنجی تخمین می‌زنند، ممکن است آن‌ها به‌اندازه کافی مسیرهای ویژگی پس‌زمینه، به‌ویژه در وجود وسایل نقلیه بزرگ در حال حرکت یا اختلاف منظر قوی را به دست نیاورند و نمی‌توانند یک هموگرافی دقیق را به دست آورند. برای تثبیت فیلم‌ها برای حل این مسئله، این مقاله روشی جدید را ارائه می‌دهد که از مسیرهای ویژگی پیش‌زمینه برای کمک به مسیرهای ویژگی پس‌زمینه در تثبیت فیلم‌های راهنمایی استفاده می‌کند. به‌طور خاص، نماهای تثبیت شده از مسیرهای پیش‌زمینه از طریق حل یک بهینه‌سازی مکانی-مکانی به دست می‌آید. سپس، مسیرهای ویژگی پس‌زمینه، به همراه مسیرهای ویژگی پیش‌زمینه و نماهای تثبیت شده آن‌ها، برای محاسبه هموگرافی موردنظر که برای تثبیت فیلم‌ها اجرا شده است، استفاده می‌شوند. برای بهبود بیشتر عملکرد تثبیت، این مقاله یک استراتژی مبتنی بر بلوک را پیشنهاد می‌کند که هر فریم را به چند بلوک مساوی تقسیم می‌کند و تلاش می‌کند تعادل تعداد صفات ویژگی بلوک‌ها را با استخراج یکنواخت نقاط برجسته و جبران مسیرهای ویژگی ناپیوسته در صورت حفظ لازم است نتایج تجربی تأیید می‌کند که روش پیشنهادی در تثبیت فیلم‌های راهنمایی و راندگی، به‌ویژه در مواردی که از اشیاء بزرگ در حال حرکت و یا اختلاف منظر برخوردار هستند، برتر از چندین روش برتر از هنر است.

همچنین در پژوهش‌های انجام شده در مقاله‌ای [۶۹] مورد بحث، به‌منظور انتخاب بهترین روش پایدارسازی تصویر ابتدا تکنیک‌های مختلف تثبیت‌کننده فیلم را مورد بررسی قرار داده است و در ادامه الگوریتمی را توسعه داده است که فرایند پایدارسازی تصاویر را با میزان خطای کمتری انجام می‌دهد، درواقع در این مدل از یک نسخه بهینه شده تطبیق بلوک در یک فضای جستجوی محدود استفاده شده است بدین شکل که ابتدا از فریم ورودی به‌صورت اتفاقی سه زیرتصویر استخراج می‌شود، سپس هیستوگرام هر سه زیرتصویر استخراج و زیرتصویری انتخاب می‌گردد که جزییات بیشتری داشته باشد. در این حالت فضای جست‌وجو کاهش می‌یابد و بهترین زیرتصویر جهت دنبال کردن در فریم‌های بعدی نیز بدست می‌آید. تا استفاده از منابع محاسباتی را به حداقل برساند. مهم‌ترین پارامتر مورد استفاده در این پژوهش نمودار جابجایی سراسری^۱ می‌باشد و نتایج این پژوهش نشان می‌دهد که روش پیشنهادی در نتایج ویدیوی خروجی مطلوب

¹ GMV

عمل کرده و در مقایسه با روش‌های دیگر به ازای پارامتر ذکر شده نیز دارای بهترین عملکرد می‌باشد و در بهترین حالت میانگین مربعات خطا این مدل برابر با ۰.۱۳۳ و ۰.۳۲۸ به ترتیب در محورهای X و Y می‌باشد.

۲-۷- جمع‌بندی

در این فصل ابتدا به بررسی پردازش تصویر پرداختیم. در ادامه نویز و دلیل وجود آن را بررسی کردیم و این بررسی را به لرزش و تاری تعمیم دادیم. در ادامه به بررسی الگوریتم‌های کلی تصحیح تصویر و پایدارسازی پرداختیم. در این بخش روش‌های حذف نویز مانند وینر، میانه و گوسی و SIFT را بیان و در انتهای فصل به بررسی سوابق کار از جمله تشریح مسئله و تخمین حرکت عمدی دوربین پرداختیم، انتهای فصل با بررسی کارهای مشابه پیشین پایان پذیرفت. در فصل بعد به معرفی روش پیشنهادی می‌پردازیم.

فصل ۳- روش پیشنهادی

۳-۱- مقدمه

همان‌طور که گفته شد به‌طور کلی وجود نویز در تصویر باعث کاهش کیفیت آن می‌شود، و تحلیل و استفاده از آن را دچار مشکل می‌کند. روش‌های رفع نویز، بخصوص در مواجهه با نویزهای قوی در تصویر، غالباً خرابی‌هایی نیز به همراه دارند. با توجه به آنکه حساسیت سیستم بینایی انسان به همه نواحی تصویر از قبیل نواحی هموار و غیر هموار به یک میزان نیست، به همین دلیل پژوهشگران به دنبال روش‌هایی هستند که میزان نویزهای موجود در تصاویر را کاهش دهد. نویزها انواع مختلفی دارند. در حوزه مطالعاتی پیش رو این نویزها لرزش‌هایی هستند که به دلایل مختلف به فریم‌ها لطمه می‌زند. نویزها از لحاظ ظاهری آزاردهنده هستند و انجام پردازش‌ها و استفاده‌های بعدی را با مشکل مواجه می‌کند، به همین دلیل یکی از مراحل اساسی در پردازش تصویر حذف نویز برای آماده‌سازی جهت پردازش‌های ثانویه بر روی آن است.

۳-۲- روش پیشنهادی

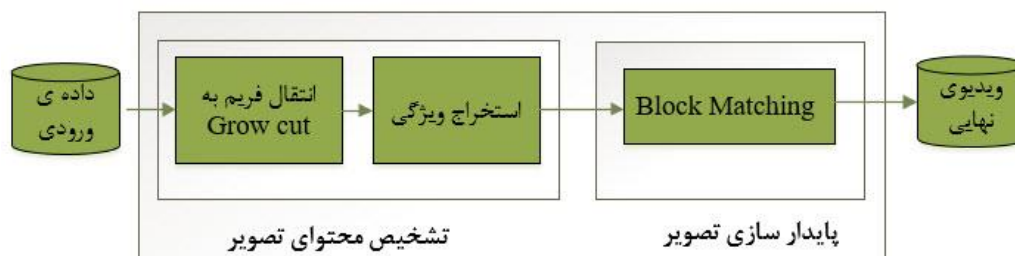
معمولاً حرکت متزلزل یک دوربین دستی بی‌ثبات یا یک حمل دوربین توسط انسان یا وجود مسیرهای ناهموار، باعث به وجود آمدن نویز در فیلم ضبط شده می‌شود که این اتفاق حس آزاردهنده‌ای را برای بیننده به وجود می‌آورد، به همین دلیل نیازمند استفاده از ابزارهای پیشرفته‌تری از قبیل ابزار تثبیت‌کننده فیلم، از بین برنده‌ی اختلالات ناخواسته ویدیو به دلیل حرکات ناپایدار دوربین هستیم.

با پیشرفت تکنولوژی بسیاری از دوربین‌ها مجهز به تثبیت‌کننده‌های سخت‌افزاری هستند، دوربین‌های گران قیمت معمولاً دارای حس‌گرها و سیستم لنزهای پیشرفته‌ای هستند تا لرزش دوربین را هنگام ضبط از بین ببرند؛ و دوربین‌های ارزان‌تر ممکن است از حس‌گرها برای تخمین حرکت دوربین و سیستم عامل داخلی برای برش دادن صحیح تصاویر به‌دست آمده استفاده کنند.

چنین راه‌حل‌های آنلاین اغلب برای برخورد با توالی‌هایی با حرکات پیچیده کافی نیستند؛ بنابراین، حتی با استفاده از دوربینی با تثبیت‌کننده آنلاین، تکان‌های ناخواسته دوربین همچنان در فیلم‌های گرفته شده توسط کاربران غیرحرفه‌ای به‌کرات اتفاق می‌افتد؛ بنابراین، برای ثابت نگه‌داشتن این فیلم‌ها هنوز الگوریتم‌های تثبیت‌کننده ویدیوی آفلاین لازم است.

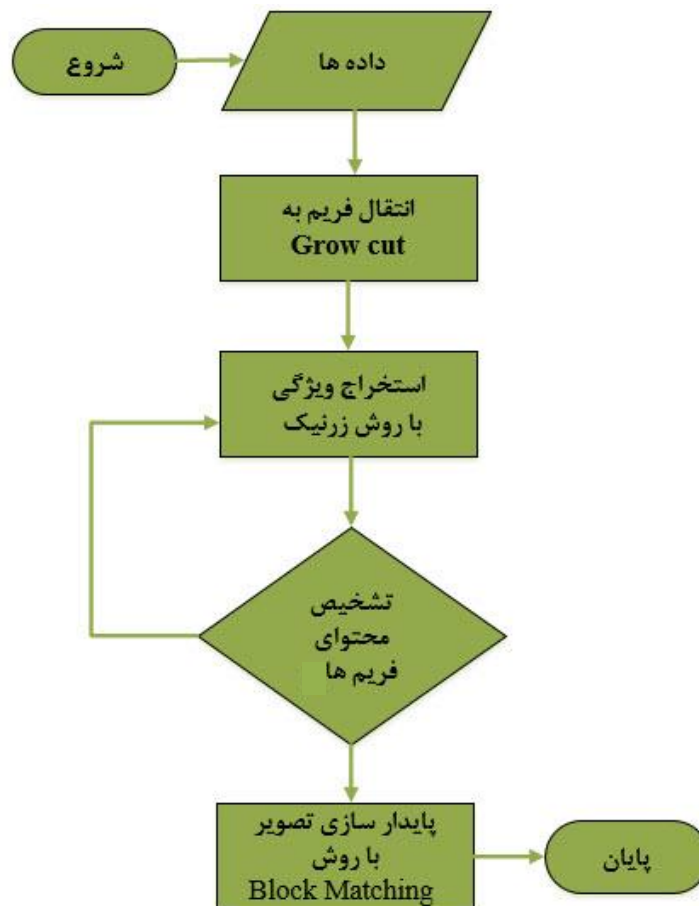
امروزه به منظور پایداری سازی ویدیوی آفلاین از روش های مختلفی استفاده می شود و افزایش میزان دقت در پایداری سازی ویدیوی آفلاین و همچنین کاهش میزان خطای مدل یکی از مهم ترین چالش ها در این حوزه می باشد. به همین دلیل در این پژوهش نیز از یک روش ترکیبی به منظور پایداری سازی تصویر استفاده شده است.

درواقع همان طور که در شکل ۱-۳ مشاهده می کنید، ابتدا فرایند تشخیص محتوای تصویر با استفاده از روش Grow Cut و الگوریتم زرنیک انجام می شود و همچنین در ادامه به منظور پایداری سازی تصویر از یک روش تطبیق بلوکی استفاده شده است.



شکل ۱-۳ روش پایداری سازی تصویر در مدل پیشنهادی

همان طور که در شکل ۲-۳ مشاهده می کنید، درواقع فرایند پایداری سازی تصویر در روش پیشنهادی شامل دو بخش تشخیص محتوای تصویر و پایداری سازی تصویر می باشد.



شکل ۳-۲ فلوجارت روش پیشنهادی

۳-۲-۱- انجام فرایند تشخیص محتوای تصویر

انجام فرایند تشخیص محتوای تصویر شامل دو بخش، انتقال فریم به Grow Cut و استخراج ویژگی با استفاده از الگوریتم زرنیک می باشد که در ادامه به بررسی کامل این بخش می پردازیم.

۳-۲-۱-۱- نحوه پردازش تصاویر با استفاده از الگوریتم Grow Cut

یک الگوریتم تقسیم تعاملی است که از روش سلولی اتوماتیک به عنوان یک مدل برای قطعه بندی تصویر استفاده می کند در این روش هر سلول دارای برچسب است که به آن بذر^۱ اولیه گفته می شود و در طی تکامل اتوماتیک، برخی از سلول ها همسایگان خود را گرفته و جایگزین برچسب آن ها می شود، در این الگوریتم ابتدا به شکل تصادفی به تعدادی از پیکسل های اولیه تصویر برچسبی تعلق می گیرد. حد آستانه

¹ Seed

تعریف شده باعث می‌شود هر پیکسل در مقایسه با پیکسل‌های اطراف یا به آن تعلق گیرد و یا برچسب خود را حفظ کند.

نحوه عملکرد این الگوریتم به صورت یک ماشین سلولی قطعی دوجهته یک سه‌تایی $A = (S, N, \delta)$ است که S یک مجموعه حالت غیرتهی و N یک سیستم همسایگی و $\delta: S^N \rightarrow S$ تابع انتقال محلی است. این تابع قوانین محاسبه‌ی حالت سلول‌ها در زمان $t+1$ است و حالات سلول‌های مجاور در زمان گذشته‌ی t را بدست می‌دهد. یک تصویر دیجیتال استخراج شده از محتوای ویدئو یک آرایه‌ی دوبعدی از $K \times m$ پیکسل است. یک تصویر غیربرچسب‌دار می‌تواند به‌عنوان یک حالت پیکربندی خاص از ماشین سلولی در نظر گرفته شود که فضای سلولی P بوسیله‌ی یک مجموعه آرایه با $K \times m$ تصویر و حالات اولیه‌ی $\forall p \in P$ به‌صورت زیر خواهد بود:

$$l_p = 0, \theta_p = 0, \vec{C}_p = RGB_p \quad \text{فرمول (۱-۳)}$$

که در آن RGB_p بردار سه‌بعدی رنگ p امین پیکسل در فضای RGB است. هدف نهایی از قطعه‌بندی تخصیص هر پیکسل به یکی از K برچسب خواهد بود. زمانی که کاربر فرآیند قطعه‌بندی را با توجه به پیکسل‌های تصویر موردنظر آغاز می‌کند، برچسب‌های سلول‌های موردنظر به‌ترتیب داده خواهند شد. این فرآیند حالت اولیه‌ی ماشین سلولی می‌باشد.

در مرحله‌ی $t+1$ برچسب سلول l_p^{t+1} با نیروی θ_p^{t+1} به‌وسیله‌ی قانون زیر تعیین خواهد شد:

که در آن g یک تابع کاهشی یکنواخت در بازه‌ی $[0,1]$ است که به صورت زیر محاسبه خواهد شد. [۸۰]

$$g(x) = 1 - \frac{x}{\max ||\vec{C}||_2} \quad \text{فرمول (۲-۳)}$$

می‌توان فرآیند برچسب‌گذاری پیکسل‌ها را به رشد و غلبه کردن K نوع باکتری تشبیه کرد. باکتری‌ها از پیکسل‌های هدف شروع به رشد می‌کنند و سعی بر اشغال کل تصویر دارند. به‌همین دلیل است که این را GrowCut نامیده‌اند. قوانین رشد رقابت بین باکتری‌ها واضح است. در هر مرحله‌ی جدا هر سلول سعی دارد تا به سلول‌های مجاور خود حمله کند. نیروی حمله توسط نیروی حمله‌کننده‌ی سلول θ_q تعیین خواهد شد که همان فاصله‌ی بین بردار ویژگی حمله‌کننده ($\vec{C}_q$) و بردار ویژگی دفاع‌کننده ($\vec{C}_p$) می‌باشد و X برابر

$\|\vec{C_p} - \vec{C_q}\|$ می‌باشد. اگر نیروی حمله‌کننده بیشتر از نیروی مدافع باشد سلول دفاع‌کننده شکست می‌خورد و برچسب آن تغییر خواهد کرد. نتیجه‌ی این رقابت محلی این است که قوی‌ترین باکتری سلول‌های اطراف خود را اشغال می‌کند و با موفقیت در سطح تصویر پراکنده خواهد شد. این الگوریتم بدلیل نحوه عملکرد به تکثیر باکتریایی نیز تشبیه می‌شود. [۷۰]



شکل ۳-۳ مراحل انجام فرایند تقسیم تعاملی در الگوریتم Grow Cut

همان‌طور که در شکل ۳-۳ قابل مشاهده است، این الگوریتم محتوای پیکسل‌های اولیه را با محتوای پیکسل‌های مجاور آن مقایسه می‌کند و در صورتی که پیکسل‌های مجاور با پیکسل اولیه شباهت داشته یا مشابه باشد (بر اساس حد آستانه مدنظر) آنگاه برچسب پیکسل اولیه بر روی پیکسل‌های مجاور نیز قرار می‌گیرد. این فرایند تا زمان قطعه‌بندی کل تصویر ادامه می‌یابد. همانطور که انتظار می‌رود با رسیدن به مرز هر قطعه مقایسه انجام شده از آستانه کمتر شده و قطعات یکی یکی شکل می‌گیرند.

۳-۲-۱-۲- انجام فرایند استخراج ویژگی با استفاده از الگوریتم زرنیک

الگوریتم زرنیک یک الگوریتم استخراج ویژگی است که ورودی خود را از الگوریتم Grow Cut دریافت می‌کند.

همچنین از مهم‌ترین ویژگی‌های این الگوریتم می‌توان به موارد زیر اشاره کرد: [۷۱]

۱- مقاومت نسبت به چرخش

بدین معنا که اگر سطحی در یک فریم وجود داشته باشد و در فریم بعدی همان سطح با اندکی چرخش ثبت شده باشد. الگوریتم زرنیک ویژگی‌های آن فریم را به‌عنوان ویژگی‌های جدید مورد پردازش قرار نمی‌دهد.

۲-مقاومت در برابر نویز و تغییرات جزئی در شکل

تغییر مقیاس و اندازه‌ی تصویر تأثیری بر ویژگی‌های استخراج شده از هر فریم ندارد.

۳-ارایه توصیف‌گرهای بهتر برای تصویر توسط ویژگی‌های استخراجی الگوریتم زرنیک

۴-تشخیص فرم کلی تا جزییات استخراجی از الگوهای داخل تصویر

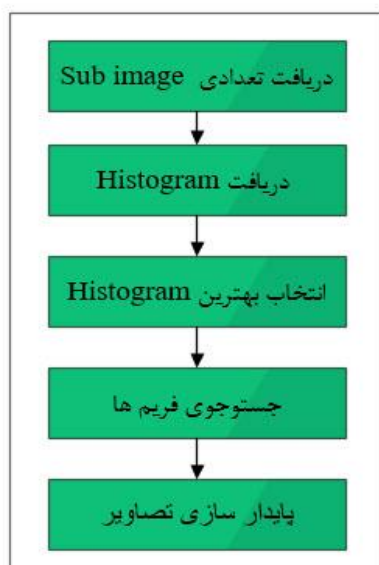
این خصوصیت‌ها باعث می‌شود ویژگی‌های مناسب از تصاویر استخراج شوند و در هر گام ویژگی‌های فریم با فریم بعدی مقایسه شوند. در نتیجه شناسایی ویژگی‌های مربوط به هر فریم دقیق‌تر و به موجب آن دسته‌بندی فریم‌ها در فرآیند تشخیص محتوا به درستی صورت می‌پذیرد. این مرحله براساس داده‌های حاصل از قطعه‌بندی که توسط Grow Cut انجام می‌گیرد، صورت گرفته و تغییر محتوا را با توجه به تغییر در ویژگی‌های استخراج شده تشخیص می‌دهد.

همان‌طور که گفته شد در بخش مربوط به تشخیص محتوای صحنه خروجی حاصل از الگوریتم زرنیک فریم‌های صحنه‌های متفاوت را از هم تفکیک می‌کند.

۳-۲-۲- فرایند پایدارسازی تصویر

در این قسمت به‌منظور پایدارسازی تصویر از یک روش تطبیق بلوکی استفاده شده است، همان‌طور که در

شکل ۳-۴ مشاهده می‌کنید، روش پایدارسازی پیشنهادی شامل پنج مرحله‌ی کلی می‌باشد. [۶۹]



شکل ۳-۴ مراحل پایدارسازی تصویر در روش تطبیق بلوکی پیشنهادی

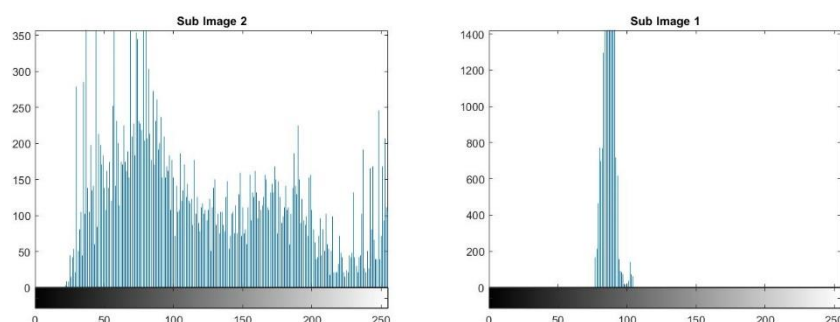
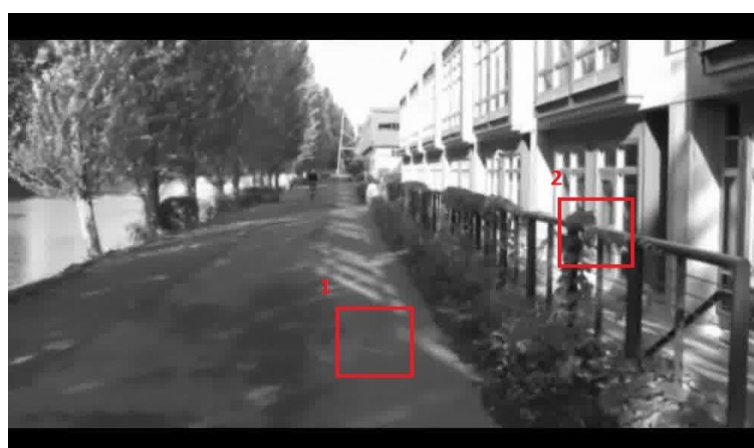
۱- دریافت تصادفی تعدادی زیرتصویر^۱ از فریم‌های تصویر

۲- گرفتن هیستوگرام زیرتصویرهای مدنظر

۳- انتخاب زیرتصویر دارای هیستوگرام با بیشترین جزئیات

۴- جست‌وجوی زیرتصویر انتخابی در فریم‌های بعدی

در این مرحله روش پیشنهادی بر اساس زیرتصویر انتخابی در فریم‌های بعدی به دنبال این زیرتصویر می‌گردد و با پیدا کردن آن در فریم‌های بعدی میزان جابه‌جایی آن را تشخیص می‌دهد و آن را جایگذاری می‌کند. برای مثال هیستوگرام دو زیرتصویر از یکی از فریم‌های ویدیو در شکل ۳-۵ آمده همان‌طور که مشهود است زیرتصویر دوم چون دارای هیستوگرام با جزئیات بیشتر است انتخاب می‌گردد.

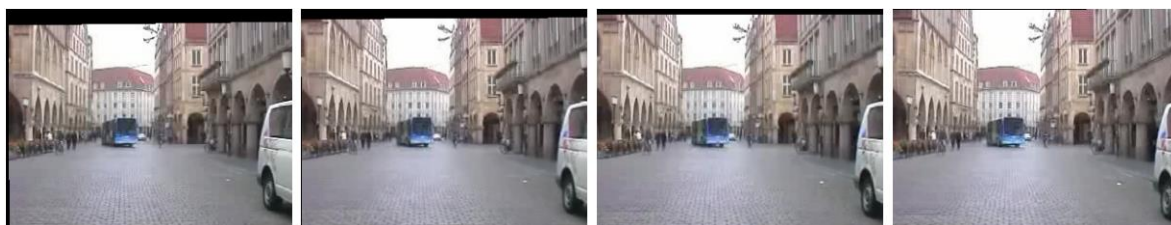


شکل ۳-۵ هیستوگرام زیرتصویرهای انتخابی

¹ Sub Image

۶- پایداری سازی تصویر

پس از حصول نتیجه مرحله قبل نمودار جابجایی سراسری بدست می‌آید. این نمودار (سیگنال) با استفاده از یک فیلتر پایین‌گذر هموار شده تا مقادیر جابجایی عمدی از غیرعمدی حذف شود و با مقادیر GMV حاصل فریم‌ها مختصات جابجایی جدید جهت پایداری سازی را پیدا کرده (مقادیر در محور X و محور Y) و مرحله ساخت فریم جبران شده یا تصحیح جابجایی^۱ انجام می‌گیرد. در این مرحله خروجی نهایی مدل مطابق شکل ۳-۶ می‌باشد. درواقع این فرایند تا انتهای فریم‌های موجود در ویدیو انجام می‌شود و بر اساس آن تصویر را پایدار می‌کند.



شکل ۳-۶ نمونه ای از فرایند پایداری سازی تصویر [۱]

۳-۳- جمع‌بندی

در این فصل ابتدا به توضیح کامل روش پیشنهادی پرداختیم و همان‌طور که گفته شد، در این پژوهش ابتدا فرایند تشخیص محتوای تصویر با استفاده از روش $Grow Cut$ و الگوریتم زرنیک انجام می‌شود و همچنین در ادامه به‌منظور پایداری سازی تصویر از یک روش تطبیق بلوک استفاده شده است، در ادامه در فصل چهارم به بررسی نتایج و خروجی مدل پیشنهادی می‌پردازیم سپس خروجی این مدل را با مدل‌های دیگر مقایسه خواهیم کرد.

¹ Motion Correction

فصل ۴- پیاده‌سازی مدل پیشنهادی

۴-۱- مقدمه

همان‌طور که گفته شده روش‌های تثبیت‌کننده تصویر فرایندهای شناخته شده‌ای هستند که برای کاهش لرزش یا جابجایی در توالی فریم‌ها استفاده می‌شود که به دلیل لرزش دوربین اتفاق می‌افتد و منجر به جلوه‌های بصری ناخوشایند در توالی‌های ویدئویی می‌شوند.

سیستم‌های تثبیت‌کننده تصویر نوری برای کاهش حرکت ظاهری در توالی تصویر با کنترل مسیر نوری حس شده توسط حس‌گرهایی مانند ژيروسکوپ یا شتاب سنج‌ها طراحی شده‌اند، درواقع این سیستم‌ها با تغییر عدسی با استفاده از دستگاه‌های اپتومکترونیکی، مسیر نوری خود را تغییر می‌دهند، همچنین در روش‌های مبتنی بر تثبیت‌کننده تصویر دیجیتالی از الگوریتم‌های پردازش تصویر به منظور پایداریسازی تصاویر استفاده می‌شود، در ادامه در این فصل ابتدا نتایج روش پیشنهادی را مورد ارزیابی قرار می‌دهیم سپس این نتایج را با پژوهش‌های دیگر مقایسه می‌کنیم.

۴-۲- پیاده‌سازی مدل پیشنهادی

در این پژوهش به منظور پیاده‌سازی مدل پیشنهادی از محیط برنامه‌نویسی متلب و همچنین روش Grow Cut استفاده شده است، همان‌طور که گفته شد در این مدل ابتدا فرایند تشخیص محتوای تصویر با استفاده از روش Grow Cut و الگوریتم زرنیک انجام می‌شود و همچنین در ادامه به منظور پایداریسازی تصویر از روش تطبیق بلوکی استفاده شده است.

۴-۲-۱- داده‌های مدل پیشنهادی

بدلیل این که در زمینه‌ی پایداریسازی ویدئو مجموعه داده‌ای استاندارد و جامع وجود ندارد، همچنین نیاز این تحقیق به ویدیویی دارای شرایط خاص که شامل برش و تغییر صحنه در فریم‌ها باشد، در این پژوهش از یک ویدیویی شخصی‌سازی شده با اتصال دو ویدئو برای ارزیابی مدل پیشنهادی استفاده شده است و همچنین به منظور مقایسه‌ی نتایج مدل پیشنهادی با مدل ارائه شده در مقاله‌ی پایه نیز از همین ویدئو

استفاده خواهد شد، درواقع به منظور ایجاد شرایط یکسان در این پژوهش، مدل ارائه شده در مقاله‌ی پایه نیز با استفاده از همین ویدیویی مورد ارزیابی قرار گرفته است.

ویدیو ورودی حاصل با نرخ برورسانی ۳۰ فریم در ثانیه بوده و حرکت آن شامل جابجایی در محور افقی و عمودی (X,Y) می‌باشد.

نکته شایان توجه جهت مقایسه ویدیوی خروجی با ورودی، عدم تغییر سرعت تصویر بوده و همچنین سرعت پخش ویدیو در جابجایی تاثیر نخواهد داشت و جابجایی بدست آمده صرفاً ناشی از لرزش فریم می‌باشد.

۴-۲-۲- پارامترهای ارزیابی مدل پیشنهادی

همان‌طور که در بخش (۱-۲) گفته شد از چالش‌های پیش‌رو در انجام این تحقیق نبود پارامترهای آماری و مشخص جهت ارزیابی روش‌های تثبیت‌سازی فریم‌هاست و بیشتر این ارزیابی به شکل عینی و از طریق سنجش میزان پایداری توسط ناظر انسانی صورت می‌گیرد ولی در کنار این مساله سعی شده به منظور بررسی و ارزیابی نتایج مدل در این پژوهش از پارامترهای ذکر شده استفاده شود تا در کنار نتایج ویدیویی حاصل بتوان سنجش مناسبی از میزان بهبودی عملکرد این الگوریتم داشت. در ادامه به توضیح کامل این پارامترها می‌پردازیم.

• نمودار حرکت سراسری (GMV):

این پارامتر نمودار میزان جابجایی هر فریم در محور Y,X را مشخص می‌کند و نمودار مربوط به آن در ادامه توضیح داده خواهد شد و متغیر K در این پارامتر یک مقدار بهینه بین ۰.۸۷۵ و ۰.۹۹۵ و تعیین کننده میزان جذب لرزش یا از دست رفتن فضای فریم و وجود فضای خالی می‌باشد. همچنین مقادیر GMV برای محاسبه مقدار میانگین جابجایی در هر دو محور در ادامه برای ارزیابی مدل استفاده خواهد شد.

$$GMV_{integrated}(t) = k \times GMV_{integrated}(t-1) + GMV_{estimated}(t-1) \quad \text{فرمول (۴-۱)}$$

- پارامتر واریانس:

واریانس عددی است که نشان می‌دهد چگونه یک سری داده حول مقدار میانگین پخش می‌شوند، برای تعریف واریانس اگر فرض کنیم که متغیر تکی X دارای توزیع $P(X)$ است و متوسط توزیع جمعیت آن را با μ نشان دهیم آنگاه واریانس این جمعیت به صورت زیر تعیین می‌شود:

$$\text{فرمول (۲-۴)} \quad \text{Var}(X) = \sigma^2 = (X - \mu)^2$$

ابتدا داده‌های ورودی نرمال شده و سپس مقدار واریانس فریم‌های انتخابی در هر دو محور X و Y محاسبه می‌شود. این مقدار بیانگر نحوه پراکندگی داده‌ها حول میانگین بوده و بنابراین میزان صحت و قابلیت اطمینان مقادیر حاصل را نشان می‌دهد. مقادیر کمتر واریانس نشان‌دهنده تجمع تعداد بیشتری از داده‌ها حول میانگین آن‌هاست.

- میانگین مربعات جابجایی:

این پارامتر به منظور مقایسه نتایج حاصل از ویدیوی ورودی، روش پایه و روش پیشنهادی تعریف شده است.

$$\text{فرمول (۳-۴)} \quad \frac{1}{n} \sum_{i=1}^n (Y_i - Y_j)^2$$

که با محاسبه میانگین مربعات میزان جابجایی حاصل از نمودار حرکت سراسری به ازای هر فریم در ویدیوی ورودی و فریم نظیر آن، هم در ویدیوی حاصل از روش پایه و هم روش پیشنهادی و برای هر دو محور X و Y بدست می‌آید.

به طور مثال در محور X میزان جابجایی در ویدیوی ورودی برای فریم n محاسبه می‌شود، این جابجایی در دو ویدیوی حاصل از روش‌های مورد بحث محاسبه شده و هرچه فاصله بین این دو مقدار بیشتر باشد به معنای جابجایی کمتر فریم و در نتیجه تثبیت بیشتر توالی فریم‌ها می‌باشد.

۴-۳- نتایج مدل پیشنهادی

در این قسمت نتایج مدل پیشنهادی را در دو بخش کلی مورد ارزیابی قرار می‌دهیم.

بخش اول: ارزیابی نتایج مدل پیشنهادی بر اساس پارامترهای واریانس و میانگین مربعات جابجایی

بخش دوم: مقایسه‌ی نتایج مدل پیشنهادی با نتایج مقاله‌ی پایه.

۴-۳-۱- ارزیابی نتایج مدل پیشنهادی بر اساس پارامترهای واریانس و

میانگین مربعات جابجایی

همان‌طور که گفته شد، در این قسمت به‌منظور ارزیابی نتایج مدل پیشنهادی تعدادی از فریم‌های یک ویدیو به‌عنوان نمونه‌های آزمایشی انتخاب شده است.

جدول ۴-۱ مقایسه‌ی میزان واریانس مدل پیشنهادی با واریانس فریم‌های ورودی

نوع	محور X	محور Y
مدل پیشنهادی	۰,۰۵۱۶	۰,۰۴۵۵
داده‌های ورودی	۰,۰۸۶۹	۰,۰۷۰۵

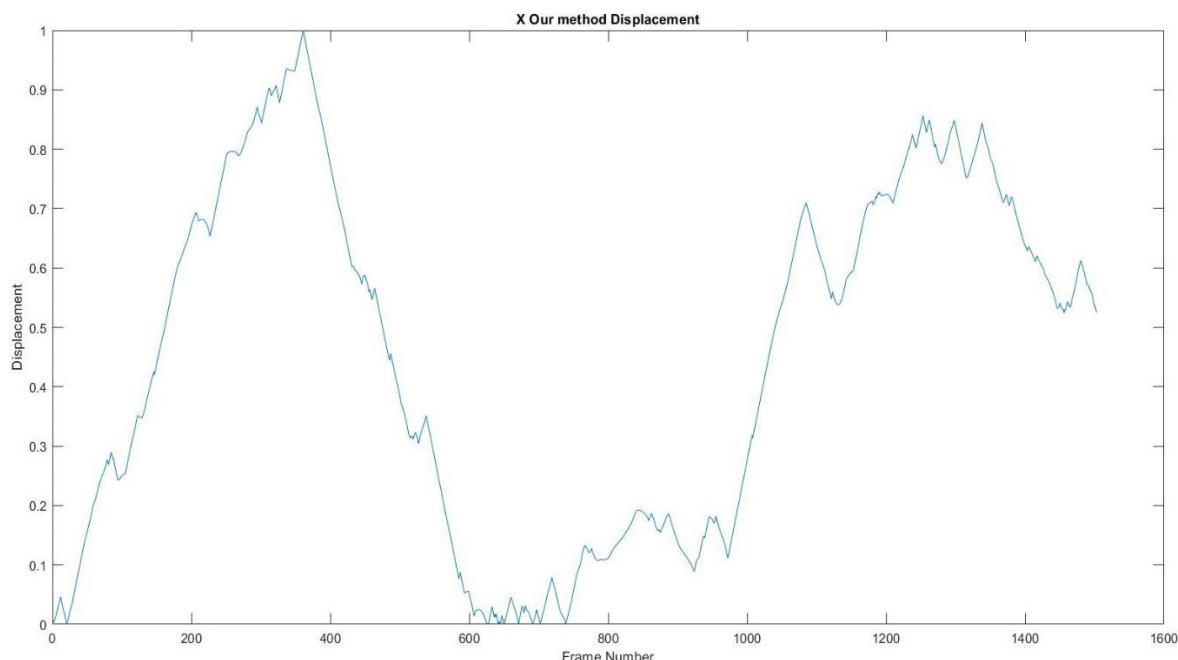
طبق نتایجی که در جدول ۴-۱ نشان داده شده است، در این قسمت ابتدا داده‌های ورودی نرمال شده و سپس مقدار واریانس فریم‌های انتخابی در هر دو محور X و Y محاسبه شده است. این مقدار بیانگر نحوه پراکندگی داده‌ها حول میانگین بوده و بنابراین میزان صحت و قابلیت اطمینان مقادیر حاصل را نشان می‌دهد. میزان واریانس مدل در محور X برابر با ۰,۰۵۱۶ و میزان واریانس مدل در محور Y نیز برابر با ۰,۰۴۵۵ می‌باشد. که به معنای تجمع اکثر داده‌ها حول میانگین می‌باشد.

یکی دیگر از پارامترهای ارزیابی عددی قابل تعریف میانگین مقادیر جابجایی یا حرکت در راستای محور X و Y می‌باشد. که نتایج آن برای روش ارایه شده در جدول زیر آمده است.

جدول ۴-۲ محاسبه‌ی میانگین جابجایی تصویر در محور X و Y بر اساس ارقام GMV

نوع	مقدار GMV در محور X	مقدار GMV در محور Y
داده‌های ورودی	۰,۵۲۶۲	۰,۴۹۷۳
مدل پیشنهادی	۰,۴۵۰۹	۰,۴۱۶۱

جدول ۲-۴ مربوط به میزان جابجایی تصویر در محورهای X و Y بر اساس ارقام GMV که این مقدار باید کمینه باشد و کم بودن این مقدار بیانگر حرکت کمتر و نرم‌تر تصویر می‌باشد. همان‌طور که در بالا بیان شد مقادیر GMV در محور X در مدل پیشنهادی برابر با ۰۰۴۵۰۹ و در محور Y برابر با ۰۰۴۱۶۱ می‌باشد.



شکل ۴-۱ میزان جابجایی تصویر در محور X بر اساس ارقام GMV در مدل پیشنهادی

شکل ۴-۱ مربوط به میزان جابجایی سراسری تصویر در محور X برای فریم‌های ویدیوی مورد مطالعه می‌باشد و همان‌طور که مشاهده می‌کنید مقدار GMV در محور X در ابتدا از فریم (۰ تا ۵۰۰) بسیار زیاد است و در ادامه در فریم‌های (۵۰۰ تا ۱۰۰۰) مقدار GMV به‌صورت چشمگیری کاهش یافته است و همچنین در فریم‌های انتهایی میزان جابجایی تصویر در محورهای X و Y روبه کاهش بوده است.

	فریم شماره‌ی ۱۲	فریم شماره‌ی ۱۴
فریم ورودی		
خروجی مدل پیشنهادی		

شکل ۲-۴ مقایسه‌ی خروجی مدل پیشنهادی با فریم‌های ورودی

شکل ۲-۴ مربوط به مقایسه‌ی تصاویر ورودی با خروجی مدل پیشنهادی می‌باشد و همان‌طور که مشاهده می‌کنید عملکرد مدل پیشنهادی در رابطه با میزان پایدارسازی تصویر نسبت به فریم‌های ورودی بهتر می‌باشد.

۴-۳-۲- مقایسه‌ی نتایج مدل پیشنهادی با نتایج مدل ارائه شده در

مقاله‌ی پایه در شرایط یکسان

در این قسمت ابتدا عملکرد مدل پیشنهادی را بر اساس داده‌های نمودار جابجایی سراسری، پارامترهای میانگین، واریانس و میانگین مربعات جابجایی با مدل ارائه شده در مقاله‌ی پایه [۶۹]، که روش مبتنی بر تطبیق بلوک است مورد مقایسه قرار می‌دهیم. و در انتهای این قسمت تعدادی از فریم‌های مربوط به خروجی مدل پیشنهادی را با فریم‌های خروجی مدل ارائه شده در مقاله‌ی پایه مقایسه می‌کنیم.

جدول ۳-۴ مقایسه‌ی میزان واریانس مدل مقاله‌ی پایه [۶۹]، با فریم‌های ورودی

نوع	محور X	محور Y
داده‌های ورودی	۰,۰۸۶۹	۰,۰۷۷۲
مدل مقاله‌ی پایه	۰,۰۷۴۵	۰,۰۷۰۵
مدل پیشنهادی	۰,۰۵۱۶	۰,۰۴۵۵

همان‌طور که در جدول ۳-۴ مشاهده می‌کنید، در این قسمت ابتدا داده‌های فریم‌های ورودی نرمال شده است و بعد از آن مقدار واریانس فریم‌های انتخابی در هر دو محور X و Y محاسبه شده است و باتوجه به مقادیر پارامتر واریانس، واریانس برای مقادیر مدل پیشنهادی کمتر بوده است بدین معنی که مدل پیشنهادی با حفظ مقادیر حول میانگین داده‌ها جابجایی را کاهش داده است.

در این تحقیق برای درک بهتر و مقایسه دقیق‌تر نتایج از معیار میانگین مربعات جابجایی بدین شکل استفاده شده است؛ از ویدیوی ورودی به عنوان مرجع استفاده شده است. برای هر دو حالت مقایسه روش پیشنهادی و روش مقاله پایه میزان جابجایی‌ها با ویدیوی ورودی، که لرزش و جابجایی زیادتری دارد، سنجیده شده است. استفاده از میانگین مربعات جابجایی بدین شکل میزان جابجایی در دو حالت را سبب می‌شود. ۱- میانگین مربعات جابجایی بین مقادیر ورودی با روش مقاله پایه و ۲- میانگین مربعات جابجایی بین مقادیر ورودی با روش پیشنهادی.

بنابراین در این شکل استفاده از این پارامتر، مقادیر پارامتر بدست آمده هرچه بیشتر باشد گویای کمتر شدن میزان جابجایی و در نتیجه پایدار شدن بیشتر ویدیو می‌باشد.

جدول ۴-۴ مقایسه‌ی میانگین مربعات جابجایی مدل پیشنهادی و روش پایه با فریم‌های ورودی

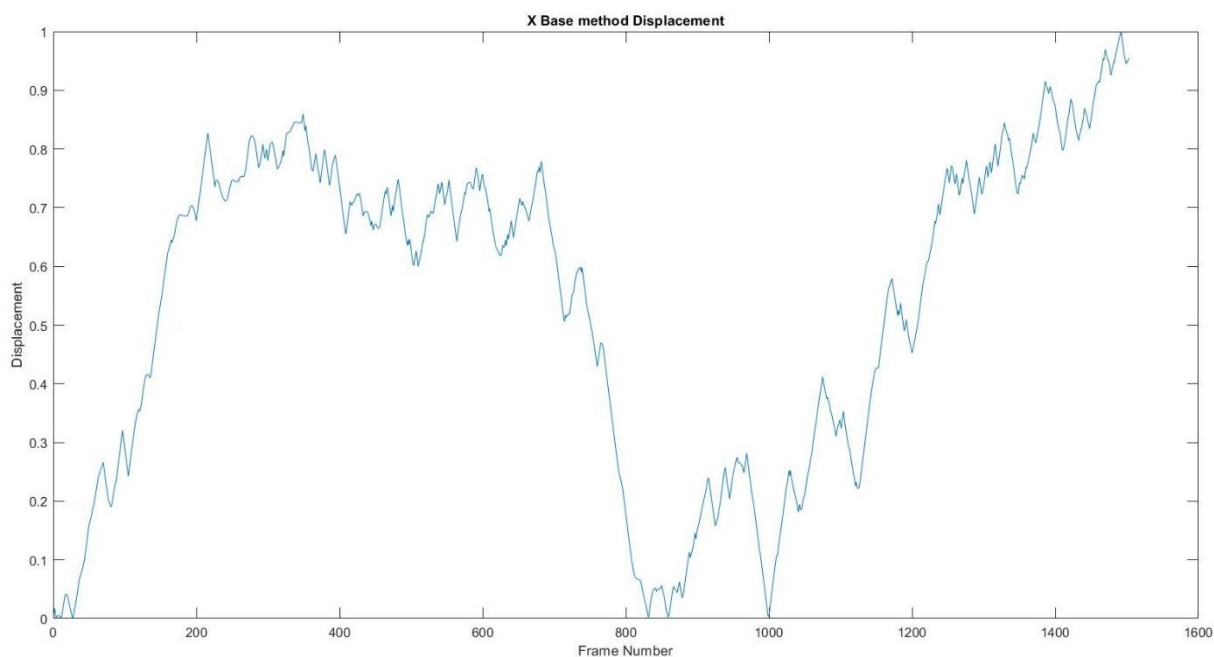
نوع	محور X	محور Y
میانگین مربعات جابجایی بین ورودی و مقاله پایه	۰,۰۷۰۸	۰,۰۸۶۷
میانگین مربعات جابجایی بین ورودی و مدل پیشنهادی	۰,۳۲۸۱	۰,۱۳۳

همان‌طور که در جدول ۴-۴ قابل مشاهده است، میزان میانگین مربعات جابجایی تعریف شده در مدل پیشنهادی افزایش چشمگیر یافته است و میزان این پارامتر در محور X برابر با ۰,۳۲۸۱ و در محور Y برابر با ۰,۱۳۳ می‌باشد که این مقدار نسبت به مدل ارائه شده در مقاله‌ی پایه بیشتر می‌باشد. همان‌طور که پیش‌تر گفته شد به دلیل سنجش جابجایی فریم‌ها به نسبت ویدیوی ورودی توسط این پارامتر، بیشتر بودن این مقدار به معنای پایداری بیشتر ویدیو می‌باشد.

جدول ۴-۵ میانگین میزان جابجایی تصویر در محور X و Y برای داده‌های ورودی، مقاله پایه و روش پیشنهادی

نوع	میانگین جابجایی در محور X	میانگین جابجایی در محور Y
داده‌های ورودی	۰,۵۲۶۲	۰,۴۹۷۳
مدل مقاله‌ی پایه	۰,۵۳۸	۰,۴۵۰۵
مدل پیشنهادی	۰,۴۵۰۹	۰,۴۱۶۱

جدول ۴-۵ مربوط به میانگین میزان جابجایی تصویر در محورهای X و Y بر اساس ارقام GMV که در بهینه‌ترین حالت مقدار GMV در محور X و Y در مدل پیشنهادی نسبت مدل مقاله‌ی پایه کمتر می‌باشد.



شکل ۴-۳ میزان جابجایی تصویر در محور X بر اساس ارقام GMV در مقاله‌ی پایه

شکل ۳-۴ مربوط به محاسبه‌ی میزان جابجایی تصویر در محور X بر اساس ارقام GMV برای فریم‌های ویدیوی انتخابی در مدل مقاله‌ی پایه می‌باشد و همان‌طور که مشاهده می‌کنید مقدار GMV از فریم (۰ تا ۲۰۰) افزایش یافته است و در ادامه از فریم (۲۰۰ تا ۷۰۰) مقدار GMV نوسانی بوده است و در ادامه در فریم‌های (۷۰۱ تا ۱۰۰۰) مقدار GMV کاهش یافته است و در ادامه با افزایش تعداد فریم‌ها میزان جابجایی تصاویر نیز افزایش یافته است.

	فریم شماره‌ی ۴۶	فریم شماره‌ی ۴۷
فریم ورودی		
خروجی مقاله پایه		

شکل ۴-۴ مقایسه‌ی خروجی مدل مقاله‌ی پایه با فریم‌های ورودی

شکل ۴-۴ مربوط به مقایسه‌ی تصاویر ورودی با خروجی مدل پیشنهادی می‌باشد و همان‌طور که مشاهده می‌کنید عملکرد مدل پیشنهادی در رابطه با میزان پایداری تصویر نسبت به فریم‌های ورودی بهتر می‌باشد.

	فریم شماره‌ی ۴۶	فریم شماره‌ی ۴۷
خروجی مدل پیشنهادی		
خروجی مقاله پایه		

شکل ۴-۵ مقایسه‌ی خروجی مدل پیشنهادی با خروجی مقاله‌ی پایه

شکل ۴-۵ مربوط به مقایسه‌ی خروجی مدل پیشنهادی و خروجی مقاله‌ی پایه می‌باشد و همان‌طور که مشاهده می‌کنید فرایند پایدارسازی تصویر در مدل پیشنهادی نسبت به مدل مقاله‌ی پایه بهتر می‌باشد.

۴-۴- نتیجه‌گیری

در این فصل ابتدا به ارزیابی نتایج مدل پیشنهادی بر اساس نمودار جابجایی سراسری و پارامترهای واریانس، میانگین جابجایی در هر دو محور X و Y و میانگین مربعات جابجایی پرداخته شده است و در ادامه به منظور بررسی دقیق‌تر خروجی مدل پیشنهادی، نتایج این پژوهش با نتایج مدل ارائه شد در مقاله‌ی پایه مقایسه شده است که در این قسمت نیز میانگین جابجایی برای روش پیشنهادی در هر دو محور و برای روش پایه در محور Y بهبود داشته است، میزان واریانس مدل پیشنهادی نسبت به مقاله‌ی پایه بهتر بوده است و همچنین میزان میانگین مربعات جابجایی معرفی شده در مدل پیشنهادی نسبت به مدل مقاله‌ی پایه افزایش محسوسی داشته است.

فصل ۵ – نتیجه‌گیری و پیشنهادها

۵-۱- نتیجه گیری

همان‌طور که گفته شد، تصاویر ثبت شده‌ی غیرحرفه‌ای معمولاً از لرزش و تکان‌های ناخواسته حس‌گر دوربین رنج می‌برند و معمولاً به دلیل حرکت متزلزل یک دوربین دستی بی‌ثبات یا یک دوربین نصب شده روی خودرو، باعث به وجود آمدن نویز در فیلم ضبط شده می‌شود که این اتفاق احساس آزاردهنده‌ای را برای بیننده به وجود می‌آورد. و هرروز به حجم این ویدیوها افزوده می‌گردد.

امروزه به‌منظور پایداری تصاویر ویدیوی از روش‌های مختلفی استفاده می‌شود و افزایش میزان دقت در پایداری تصاویر ویدیوی و همچنین کاهش میزان خطای مدل یکی از مهم‌ترین چالش‌ها در این حوزه می‌باشد، به همین دلیل در این پژوهش نیز از یک روش ترکیبی به‌منظور پایداری فریم‌های ویدیو استفاده شده است.

در این مدل ابتدا فرایند تشخیص محتوای تصویر با استفاده از روش Grow Cut و الگوریتم زرنیک انجام شده است و در ادامه به‌منظور پایداری تصاویر از یک روش تطبیق بلوکی استفاده شده است، درواقع فرایند پایداری تصویر در روش پیشنهادی شامل دو بخش تشخیص محتوای تصویر و پایداری تصویر می‌باشد. در بخش اول فرایند تشخیص محتوای تصویر انجام می‌شود، درواقع انجام فرایند تشخیص محتوای تصویر شامل دو بخش، انتقال فریم به Grow Cut برای قطعه‌بندی تصاویر و استفاده از الگوریتم زرنیک جهت استخراج ویژگی‌های مقاوم در برابر چرخش و تغییر مقیاس و تغییرات جزئی در شکل می‌باشد.

در بخش دوم فرایند پایداری تصویر انجام می‌شود، در این قسمت به‌منظور پایداری تصاویر از یک روش تطبیق بلوکی استفاده شده است. که با استخراج زیرتصویر بهینه از فریم ورودی و یافتن آن در فریم‌های بعدی بردار جابجایی سراسری را جهت حصول مقادیر جابجایی غیرعمدی برای تصحیح فریم و پایداری تصاویر انجام می‌گیرد.

از آنجا که در زمینه‌ی پایداری ویدیو مجموعه داده‌ی استاندارد وجود ندارد در این پژوهش از یک فایل ویدیویی متشکل از دو ویدیو با موضوعات متفاوت برای ارزیابی مدل پیشنهادی استفاده شده است و به‌منظور مقایسه‌ی نتایج مدل پیشنهادی با مدل ارائه شده در مقاله‌ی پایه نیز از همین ویدیو استفاده شده است، در

ادامه به منظور ارزیابی مدل پیشنهادی از پارامترهای نمودار جابجایی سراسری، میانگین جابجایی در هردو محور، واریانس در هر دو محور و میانگین مربعات جابجایی برای هردو روش پایه و پیشنهادی استفاده شده است.

نتایج این پژوهش نشان می‌دهد که مقدار واریانس مدل پیشنهادی در محور X برابر با ۰۰۰۵۱۶ و میزان واریانس مدل در محور Y نیز برابر با ۰۰۰۴۵۵ می‌باشد که در مقایسه با نتایج قبلی عملکرد بهتر روش را نشان می‌دهد، همچنین میزان جابجایی تصویر در محورهای X و Y بر اساس ارقام GMV در بهینه‌ترین حالت مقدار GMV در محور X در مدل پیشنهادی برابر با ۰۰۴۵۰۹ و در محور Y برابر با ۰۰۴۱۶۱ بوده که همچنین نمایان‌گر بهبود در روش پیشنهادی می‌باشد، در انتها نیز نتایج این مدل با مدل ارائه شده در مقاله‌ی پایه [۶۹]، مقایسه شده است که در این قسمت مدل پیشنهادی به ازای پارامترهای ذکر شده دارای عملکرد بهتری بوده است.

۵-۲- کار آینده

یادگیری عمیق دارای شبکه‌ای بسیار قدرتمند است که هوش مصنوعی به‌کار رفته در آن به مراتب قدرتمندتر از هوش انسان عمل می‌کند، پتانسیل بیشتر جهت یادگیری داشته و سرعت پاسخ‌گویی بالایی دارد. به این صورت که در آغاز ساختارهای ابتدایی داده‌های ورودی را دریافت کرده سپس با ترکیب توابع خطی و غیرخطی بسیار، ساختارهایی پیچیده و سطح بالا از داده‌های ورودی تولید کرده و پردازش می‌کند. همچنین الگوریتم‌های یادگیری عمیق به دلیل داشتن فاز آموزشی، توانایی پردازش داده‌هایی با حجم بالا را دارا هستند و با افزایش حجم داده‌ها میزان دقت این الگوریتم‌ها نیز افزایش می‌یابد. در حوزه‌ی پردازش تصویر استفاده از الگوریتم‌های یادگیری عمیق نتایج مناسبی را به همراه داشته است به همین دلیل در این پژوهش نیز به‌منظور افزایش میزان دقت در پایدارسازی تصویر و همچنین کاهش میزان خطای مدل، استفاده از الگوریتم‌های یادگیری عمیق توصیه می‌شود.

- [1] Ken-Yi Lee, Yung-Yu Chuang, Bing-Yu Chen, and Ming Ouhyoung, "Video stabilization using robust feature trajectories," 2009 IEEE 12th Int. Conf. Comput. Vis., no. Iccv, pp. 1397–1404, 2009.
- [2] J. Windau and L. Itti, "Multilayer real-time video image stabilization," in IEEE International Conference on Intelligent Robots and Systems, 2011, vol. 1, no. 213, pp. 2397–2402.
- [3] L. Zhang, Q. K. Xu, and H. Huang, "A Global Approach to Fast Video Stabilization," IEEE Trans. Circuits Syst. Video Technol., vol. 27, no. 2, pp. 225–235, 2017.
- [4] A. Walha, A. Wali, and A. M. Alimi, "Video Stabilization for Aerial Video Surveillance," AASRI Procedia, vol. 4, pp. 72–77, 2013.
- [5] M. Zerara Mr, "Projection-Based Rotation Estimation for Video Stabilization," 2017.
- [6] R. J. Ortiz Cayón, "Online Video Stabilization for UAV," 2013.
- [7] M. R. Souza and H. Pedrini, "Digital video stabilization based on adaptive camera trajectory smoothing," Eurasip J. Image Video Process., vol. 2018, no. 1, 2018.
- [8] L. Linlin, "A Survey on Video Stabilization," Computing, vol. 31, no. September, pp. 1–25, 2011.
- [9] H. Shen, Q. Pan, Y. Cheng, and Y. Yu, "Fast video stabilization algorithm for UAV," Proc. - 2009 IEEE Int. Conf. Intell. Comput. Intell. Syst. ICIS 2009, vol. 4, pp. 542–546, 2009.
- [10] A. Lim, B. Ramesh, Y. Yang, C. Xiang, and Z. Gao, "Real-Time Optical flow-based Video Stabilization for Unmanned Aerial Vehicles," J. Real-Time Image Process., pp. 1–11, Jan. 2009.
- [11] Hu, Hao, and Gerard De Haan. "Low cost robust blur estimator." In 2006 International Conference on Image Processing, pp. 617-620. IEEE, 2006.
- [12] Levin, Anat, Rob Fergus, Frédo Durand, and William T. Freeman. "Image and depth from a conventional camera with a coded aperture." ACM transactions on graphics (TOG) 26, no. 3 (2007): 70.
- [13] Shan, Qi, Jiaya Jia, and Aseem Agarwala. "High-quality motion deblurring from a single image." In ACM Transactions on Graphics (TOG), vol. 27, no. 3, p. 73. ACM, 2008.
- [14] Cho, Sunghyun, and Seungyong Lee. "Fast motion deblurring." In ACM Transactions on Graphics (TOG), vol. 28, no. 5, p. 145. ACM, 2009.

- [15] Ferzli, Rony, and Lina J. Karam. "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)." *IEEE transactions on image processing* 18, no. 4 (2009): 717-728.
- [16] Tai, Yu-Wing, Hao Du, Michael S. Brown, and Stephen Lin. "Correction of spatially varying image and video motion blur using a hybrid camera." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, no. 6 (2010): 1012-1028.
- [17] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," *Proc. The 11th European Conference on Computer Vision (ECCV)*, pp. 157-170, 2010.
- [18] Caviedes, Jorge, and Franco Oberti. "A new sharpness metric based on local kurtosis, edge and energy information." *Signal Processing: Image Communication* 19, no. 2 (2004): 147-161.
- [19] Ong, EePing, Weisi Lin, Zhongkang Lu, Xiaokang Yang, Susu Yao, Feng Pan, Lijun Jiang, and Fulvio Moschetti. "A no-reference quality metric for measuring image blur." In *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, vol. 1, pp. 469-472. Ieee, 2003.
- [20] Chern, N. K., N. P. A. Neow, and M. H. Ang Jr. "Blur Determination in the compressed domain using DCT information." In *Proc. IEEE Int. Conf. Robotics and Automation*, vol. 3, pp. 2791-96. 2001.
- [21] Nill, Norman B., and Brian Bouzas. "Objective image quality measure derived from digital image power spectra." *Optical engineering* 31, no. 4 (1992): 813-825.
- [22] Marichal, Xavier, Wei-Ying Ma, and HongJiang Zhang. "Blur determination in the compressed domain using DCT information." In *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, vol. 2, pp. 386-390. IEEE, 1999.
- [23] Marichal, X., W. Ma, and H. J. Zhang. "Practical issues in pixel-based auto focusing for Machine Vision." In *Proc. IEEE Int. Conf. Image Processing*, vol. 2, pp. 386-90. 1999.
- [24] Chiang, Ming-Chao, and Terrance E. Boulton. "Local blur estimation and super-resolution." In *cvpr*, vol. 97, pp. 821-826. 1997.
- [25] Ong, EePing, Weisi Lin, Zhongkang Lu, Xiaokang Yang, Susu Yao, Feng Pan, Lijun Jiang, and Fulvio Moschetti. "A no-reference quality metric for measuring image blur." In *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, vol. 1, pp. 469-472. Ieee, 2003.

- [26] Marziliano, Pina, Frederic Dufaux, Stefan Winkler, and Touradj Ebrahimi. "A no-reference perceptual blur metric." In Image Processing. 2002. Proceedings. 2002 International Conference on, vol. 3, pp. III-57. IEEE, 2002.
- [27] Ferzli, Rony, and Lina J. Karam. "No-reference objective wavelet based noise immune image sharpness metric." In IEEE International Conference on Image Processing 2005, vol. 1, pp. I-405. IEEE, 2005.
- [28] Bong, David BL, and Adeline SL Ng. "Full-Reference and No-reference Image Blur Assessment Based on Edge Information." International Journal on Advanced Science, Engineering and Information Technology 2, no. 1 (2012): 90-95.
- [29] Caviedes, Jorge, and Sabri Gurbuz. "No-reference sharpness metric based on local edge kurtosis." In Image Processing. 2002. Proceedings. 2002 International Conference on, vol. 3, pp. III-53. IEEE, 2002.
- [30] Bahrami, Khosro, and Alex C. Kot. "A fast approach for no-reference image sharpness assessment based on maximum local variation." IEEE Signal Processing Letters 21, no. 6 (2014): 751-755.
- [31] Fezza, Sid Ahmed, and Mohamed-Chaker Larabi. "No-reference perceptual blur metric for stereoscopic images." In 2014 International Conference on 3D Imaging (IC3D), pp. 1-8. IEEE, 2014.
- [32] Crete, Frederique, et al. "The blur effect: perception and estimation with a new no-reference perceptual blur metric." Electronic Imaging 2007. International Society for Optics and Photonics, 2007.
- [33] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision.," Robotics, vol. 81, pp. 674–679, 1981.
- [34] W. G. Aguilar and C. Angulo, "Real-Time Model-Based Video Stabilization for Microaerial Vehicles," Neural Process. Lett., vol. 43, no. 2, pp. 459–477, 2016.
- [35] J.H. Kotecha, P.M. Djuric, "Gaussian particle filtering", IEEE Trans. Signal Process., vol. 51, no. 10, pp. 2592–2601, Oct. 2003.
- [36] S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, "A tutorial on particle filters for on-line non-linear/non-gaussian Bayesian tracking", IEEE Trans. Signal Processing, vol. 50, no. 2, pp. 174-188, Feb. 2002.
- [37] Sedaghat, A., & Ebadi, H. (2015). Remote sensing image matching based on adaptive binning SIFT descriptor. IEEE transactions on geoscience and remote sensing, 53(10), 5283-5293.

- [38] Sharif, M., Khan, S., Saba, T., Raza, M., & Rehman, A. (2019, April). Improved Video Stabilization using SIFT-Log Polar Technique for Unmanned Aerial Vehicles. In *2019 International Conference on Computer and Information Sciences (ICCIS)* (pp. 1-7). IEEE.
- [39] L. Kejriwal and I. Singh, "A Hybrid Filtering Approach of Digital Video Stabilization for UAV Using Kalman and Low Pass Filter," *Procedia Comput. Sci.*, vol. 93, no. 1, pp. 359–366, Jan. 2016.
- [40] C. Wang, J. Kim, K. Byun, J. Ni, and S. Ko, "Robust digital image stabilization using the Kalman filter," *IEEE Trans. Consum. Electron.*, vol. 55, no. 1, pp. 6–14, Feb. 2009.
- [41] T. N. Shene, K. Sridharan, and N. Sudha, "Real-Time SURF-Based Video Stabilization System for an FPGA-Driven Mobile Robot," *IEEE Trans. Ind. Electron.*, vol. 63, no. 8, pp. 5012–5021, 2016.
- [42] W. G. Aguilar and C. Angulo, "Real-time video stabilization without phantom movements for micro aerial vehicles," *Eurasip J. Image Video Process.*, vol. 2014, no. 1, p. 46, Dec. 2014.
- [43] M. Grundmann, V. Kwatra, and I. Essa, "Auto-directed video stabilization with robust L1 optimal camera paths," in *CVPR 2011*, 2011, no. 1, pp. 225–232.
- [44] Y.M. Liang, H.R. Tyan, S.L. Chang, H.Y.M. Liao, S.W. Chen, "Video stabilization for a camcorder mounted on a moving vehicle", *IEEE Trans. on Vehicular Technology*, vol. 53, no. 6, pp. 1636-1648, Nov. 2004.
- [45] H.H. Chen, C.K. Liang, Y.C. Peng, H.A. Chang, "Integration of digital stabilizer with video codec for digital video cameras", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 17, no. 7, pp. 801-813, July 2007.
- [46] A. Doucet, A.M. Johansen, "A Tutorial on Particle Filtering and Smoothing: Fifteen years later", Oxford University Press, Version 1.1, pp. 4-6, Dec. 2008.
- [47] S. Erturk, "Digital image stabilization with sub-image phase correlation based global motion estimation", *IEEE Trans. on Consumer Electronics*, Nov. 2003.
- [48] Z. Pan, C. Ngo, "Selective object stabilization for home video consumers", *IEEE Trans. on Consumer Electronics*, vol. 51, no. 4, Nov. 2005.
- [49] A.A. Yeni, S. Erturk, "Sast digital image stabilization using one bit transform based sub-image motion estimation", *IEEE Trans. on Consumer Electronics*, vol. 51, no. 3, Aug. 2005.

- [50] A. Litvin, J. Konrad, W.C. Karl, "Probabilistic video stabilization using kalman filtering and mosaicking", IS&T/SPIE Symposium on Electronic Imaging, pp. 663-674, 2003.
- [51] J. Yang, D. Schonfeld, M. Mohamed, "Robust Video Stabilization Based on Particle Filter Tracking of Projected Camera Motion", IEEE Trans. on Circuits and Systems for Video Technology, vol. 19, no. 7, July 2009.
- [52] Y. Shen, P. Guturu, T. Damarla, B.P. Buckles, K.R. Namuduri, "Video Stabilization Using Principal Component Analysis and Scale Invariant Feature Transform in Particle Filter Framework", IEEE Trans. on Consumer Electronics, vol. 55, No. 3, Aug. 2009.
- [53] J. Yang, D. Schonfeld, C. Chen, M. Mohamed, "Online Video Stabilization Based On Particle Filters", IEEE International Conference on Image Processing, pp. 1545-1548, Oct. 2006.
- [54] Y.C. Peng, H.A. Chang, H.H. Chen, C.J. Kao, "Digital image stabilization and its integration with video encoder", IEEE Consumer Communications and Networking Conference, pp. 544-549, Jan. 2005.
- [55] M.K. Gullu, E. Yaman, S. Erturk, "Image Sequence Stabilization Using Furzy Adaptive Kalman Filtering", Electronics Letters, vol. 39, no. 5, pp. 429-431, Mar. 2003.
- [56] M.K. Gullu, S. Erturk, "Fuzzy image sequence stabilization", Electronics Letters , vol. 39, no. 16, pp. 1170- 1172, Aug. 2003.
- [57] S. Nikitidis, S. Zafeiriou, I. Pitas, "Camera Motion Estimation Using a Novel Online Vector Field Model in Particle Filters", IEEE Trans. on Circuits and Systems for Video Technology, vol. 18, no. 8, Aug. 2008.
- [58] S.K. Zhou, R. Chellappa, B. Moghaddam, "Visual tracking and recognition using appearance-adaptive models in particle filters", IEEE Trans. Image Process., vol. 13, no. 11, pp. 1491-1506, Nov. 2004.
- [59] Fergus, Rob, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. "Removing camera shake from a single photograph." In ACM Transactions on Graphics (TOG), vol. 25, no. 3, pp. 787-794. ACM, 2006.
- [60] Yeung, F., Levinson, S. F., & Parker, K. J. (1998). Multilevel and motion model-based ultrasonic speckle tracking algorithms. *Ultrasound in medicine & biology*, 24(3), 427-441.
- [61] Golemati, S., Stoitsis, J., & Nikita, K. S. (2007). Motion analysis of the carotid artery wall and plaque using B-mode ultrasound. *Vascular Disease Prevention*, 4(4), 296-302.

- [62] Grava, C., Gacsádi, A., & Gavrilut, I. (2009). Arterial elasticity maps obtained by using basic block-matching methods. *Journal of Electrical and Electronics Engineering*, (1), 151.
- [63] Ramnarine, K. V., Kanber, B., & Panerai, R. B. (2004). Assessing the performance of vessel wall tracking algorithms: the importance of the test phantom. In *Journal of Physics: Conference Series* (Vol. 1, No. 1, p. 199). IOP Publishing.
- [64] Yeh, Y., Wang, S. J., & Chiang, H. C. (2000, June). Digital camcorder image stabilizer based on gray-coded bit-plane block matching. In *Input/Output and Imaging Technologies II* (Vol. 4080, pp. 112-120). International Society for Optics and Photonics.
- [65] Litvin, A., Konrad, J., & Karl, W. C. (2003, May). Probabilistic video stabilization using Kalman filtering and mosaicing. In *Image and Video Communications and Processing 2003* (Vol. 5022, pp. 663-674). International Society for Optics and Photonics.
- [66] Joseph, K., Raj, A. N. J., Fan, Z., & Vidhyapathi, C. M. (2017, July). A time-efficient video stabilization algorithm based on block matching in a restricted search space. In *2017 IEEE International Conference on Real-time Computing and Robotics (RCAR)* (pp. 651-656). IEEE.
- [67] Lim, A., Ramesh, B., Yang, Y., Xiang, C., Gao, Z., & Lin, F. (2019). Real-time optical flow-based video stabilization for unmanned aerial vehicles. *Journal of Real-Time Image Processing*, 16(6), 1975-1985.
- [68] Zhao, M., & Ling, Q. (2019). A robust traffic video stabilization method assisted by foreground feature trajectories. *IEEE Access*, 7, 42921-42933.
- [69] Joseph, Kevin, et al. "A time-efficient video stabilization algorithm based on block matching in a restricted search space." 2017 IEEE International Conference on Real-time Computing and Robotics (RCAR). IEEE, 2017.
- [70] Vezhnevets, Vladimir, and Vadim Konouchine. "GrowCut: Interactive multi-label ND image segmentation by cellular automata." *proc. of Graphicon*. Vol. 1. No. 4. Citeseer, 2005.
- [71] Huang, M., Shu, H., Ma, Y., & Gong, Q. (2015). Content-based image retrieval technology using multi-feature fusion. *Optik-International Journal for Light and Electron Optics*, 126(19), 2144-2148.

Abstract:

Unprofessionally captured images usually suffer from unwanted vibrations and vibrations of the camera sensor, and usually due to the shaky motion of an unstable handheld camera or a camera mounted on the car, causes noise in the recorded video, which makes the viewer feel annoyed. Today, different methods are used to stabilize video images, and increasing the accuracy of stabilizing video images and also reducing the model error is one of the most important challenges in this area, so in this study, a combined method is used to stabilize video frames.

In this model, first the image content recognition process is performed using Grow Cut method and Zernike Algorithm and then a block matching method is used to stabilize the image. In fact, the image stabilization process in the proposed method consists of two parts: image content detection and image stabilization. In the first part, the image content recognition process is performed. In fact, the image content recognition process consists of two parts, transferring the frame to Grow Cut and extracting the feature using the Zernike algorithm. In the second part: the image stabilization process is performed. In this part, a block matching method is used to stabilize the image. In order to evaluate the proposed model, the parameters of variance, mean square displacement and Global Motion Vector(GMV) parameters are used and values of GMV those used to calculate other parameters normalized between 0-1. The results of this study show that the value of variance of the proposed model in the X axis is equal to 0.0516 and the amount of variance of the model in the Y axis is equal to 0.0455, also the amount of average image movement in the X and Y axes based on GMV figures in the optimal value The GMV on the X-axis in the proposed model is 0.4509 and for Y-axis amount of 0.4161 reported.

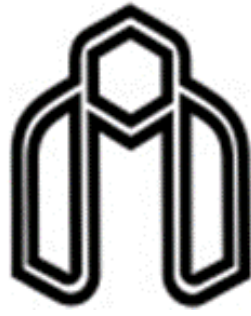
This research faced with some limits like lack of computational parameters and most of the comparison between the results should apply by human judgements from watching the final video. To overcome this problem, we used the mean square displacement measure. It represents the difference of displacement of input video with, proposed method and basic method [80].

And the results of that shows amount of 0.0708 for X-axis and 0.0867 for Y-axis between input and base method, and results equal to 0.3281 for X-axis and 0.133 for Y-axis between input and proposed method. As long as this parameter measures the displacement between input video that has most vibration and displacements, the results of the best output should

be more than the other. So it shows that the proposed method did much better in order of mean square displacement measure.

Beside that in this research there was no dataset or much used input videos, and the input video created for that by merging two input videos used by other works in this research area.

Keywords: Image Stabilization System, Block Matching Method, Feature Extraction, Zernike Algorithm



**Shahrood University of
Technology**

Faculty of Computer Engineering
MSc Thesis in Artificial Intelligence

Improving the video stabilizing performance considering the scene

By: Siavash Riyahi Sales

Supervisor:
Dr Hamid Hassanpour

Advisor:
Dr. Mohsen Biglari

March 2021