

روشی جدید در ردیابی اهداف متحرک هوایی با استفاده از الگوریتم AFFINE-SIFT

وحید آزادزاده^۱، علی محمد لطیف^{۲*}

۱- کارشناس ارشد، ۲- دانشیار، دانشگاه یزد

(دریافت: ۹۵/۰۳/۱۸، پذیرش: ۹۵/۰۸/۱۰)

چکیده

ردیابی هدف متحرک فرایندی است که در آن یک شیء مشخص در یک دنباله ویدئویی از قاب‌ها تعقیب و مکان آن در هر قاب آشکار می‌شود. هدف از این فرایند تسهیل در پردازش‌های بعدی برای تحلیل رفتار یا شناسایی سوژه متحرک است. در این مقاله رویکردی جدید در زمینه آشکارسازی و ردیابی اهداف متحرک هوایی بر مبنای الگوریتم‌های تطبیق مشخصه ارائه شده است. آشکارسازی اهداف هوایی توسط چگالی طیفی و اطلاعات پیشینه‌ی هدف انجام می‌شود و برای ردیابی اهداف متحرک از الگوریتم تطبیق مشخصه ASIFT استفاده می‌شود. چالش موجود انتخاب ویژگی‌هایی است که در مقابل تغییرهای شدت روشنایی، نویز، دوران، تغییر مقیاس و زاویه دید مقاوم باشند. برای حل این مشکل نقاط کلیدی و متناظر آن‌ها در الگوهای استخراج شده از قاب‌های متوالی، توسط الگوریتم ASIFT محاسبه می‌شود، همچنین به منظور کاهش تناظرهای اشتباه در قاب‌های متوالی از الگوریتم RANSAC استفاده شده است. در این مقاله برای مقاوم کردن الگوریتم نسبت به تغییر مقیاس هدف از تاریخچه مقیاس سوژه در ۱۰ قاب قبلی استفاده شده است. الگوریتم ارائه‌شده بر روی پایگاه داده استاندارد AIRCRAFT TRACKING اجرا شد. نتایج آزمایش‌ها کارآمدی روش ارائه شده را در دقت ردیابی نسبت به الگوریتم‌های مطرح در ردیابی نشان می‌دهد.

واژه‌های کلیدی: آشکارسازی و ردیابی هدف، الگوریتم ASIFT، الگوریتم RANSAC، تاریخچه مقیاس هدف، چگالی طیفی

۱- مقدمه

ردیابی هدف متحرک به مفهوم تعقیب هدف متحرک در تصاویر متوالی است و یکی از مسائل مهم در بینایی ماشین است. از کاربردهای ردیابی هدف متحرک می‌توان به بازشناسی ژست بر اساس حرکت (تحلیل حرکت بازیکنان، شناسایی فرد بر اساس طرز راه رفتن و روند بهبود بیماری)، نظارت خودکار صحنه‌های مشکوک (مترو، موزه و پارکینگ)، نمایه‌گذاری فایل‌های ویدئویی، ارتباط انسان و رایانه از طریق ردیابی حرکات چشم و تشخیص و ردیابی خودکار اهداف متحرک در سامانه‌های موشکی و ضد موشکی نام برد. چالش‌هایی که در ردیابی هدف متحرک وجود دارد، پوشیدگی هدف، تغییرات شدت روشنایی، تغییر حالت ظاهر هدف، صحنه شلوغ و پیچیده و حرکت دوربین است [۱].

تاکنون روش‌های مختلفی برای ردیابی اشیاء متحرک در تصاویر متوالی پیشنهاد شده است. در یک دسته‌بندی الگوریتم‌های ردیابی به چهار دسته ردیابی نقاط، ردیابی بر اساس هسته، ردیابی سیاه‌نما و روش‌های تفریق پس‌زمینه تقسیم‌بندی

می‌شوند [۲]. هر یک از الگوریتم‌های مطرح شده محدودیت‌هایی دارند که با توجه به شرایط مساله الگوریتم مناسب انتخاب می‌شود.

الگوریتم‌های ردیابی نقاط را الگوریتم‌های ردیابی مبتنی بر تطبیق مشخصه نیز می‌نامند. به دلیل ناتوانی در جداسازی پیش‌زمینه از پس‌زمینه، الگوریتم‌های ردیابی مبتنی بر تطبیق مشخصه در ردیابی اشیاء صلب نتیجه‌ی بهتری را ارائه می‌دهند. هنگامی که شیء غیرصلب باشد، ویژگی‌های استخراج شده از پس‌زمینه الگوریتم را با مشکل مواجه خواهد کرد [۳]. سرعت پردازش در روش‌های تطبیق مشخصه خوب است؛ ولی انتخاب ویژگی‌هایی که بتوانند در مقابل تغییرهای شدت روشنایی، نویز، دوران، تغییر مقیاس و زاویه دید مقاوم باشند، مشکلی است که هنوز باقی مانده است.

روش‌های ردیابی نقاط به دو دسته قطعی^۱ و آماری^۲ تقسیم‌بندی می‌شوند. روش‌های قطعی برای تطبیق نقاط با استفاده از مجموعه‌ای از شروط حرکتی، یک هزینه وابستگی

این الگوریتم توسط Lowe در سال ۱۹۹۴ ارائه شد که نسبت به تغییر مقیاس، دوران، تغییر زاویه دید و تا حدودی نسبت به تغییر شدت روشنایی مقاوم است. اثبات ریاضی مقاوم بودن SIFT در برابر تغییر مقیاس در مرجع [۸] بررسی شده است.

الگوریتم ASIFT^۴ نسخه بهبود داده شده الگوریتم SIFT است که جهت تناظریابی تصاویر، مستقل از تغییر شکل افاین ارائه شده است [۹]. ASIFT در برابر شش پارامتر تبدیلات افاین مقاوم است، در حالی که روش SIFT تنها نسبت به چهار پارامتر انتقال، دوران و تغییر مقیاس مقاوم است [۷]. به همین دلیل در الگوریتم پیشنهادی برای استخراج و تطبیق نقاط کلیدی در قاب‌های ویدئویی از الگوریتم ASIFT استفاده می‌شود.

به دلیل شبیه‌سازی تصاویر از منظرهای مختلف در الگوریتم ASIFT تعداد نقاط کلیدی استخراج شده نسبت به الگوریتم SIFT بیش‌تر است؛ لذا برای تعیین نقاط متناظر و حذف تناظرهای اشتباه از سازگاری آن‌ها با قید هندسی خط اپی‌پولار^۵ استفاده می‌شود. روش استفاده شده در الگوریتم ASIFT برای یافتن نقاط متناظر و حذف داده‌های پرت^۶، ORSA نام دارد که بهبود یافته الگوریتم RANSAC است. در روش ارائه شده در این مقاله از ترکیب دو روش ORSA و RANSAC برای یافتن نقاط تطبیقی و حذف تناظرهای اشتباه استفاده می‌شود. در شکل (۱)، ۹۳ تطبیق توسط الگوریتم ORSA به دست آمده است. پس از اعمال الگوریتم RANSAC، ۴۷ تطبیق درست شناخته شد و باقی تطابق‌ها که نادرست بودند حذف شدند.



الف) ۹۳ نقاط تطبیقی به دست آمده توسط الگوریتم ORSA



ب) ۴۷ نقاط تطبیقی استخراج شده توسط الگوریتم RANSAC
شکل (۱). یافتن نقاط تطبیقی توسط الگوریتم ORSA و RANSAC

[۱۰]

برای هر شیء در قاب $t-1$ به یک شیء در قاب t تعریف می‌کنند که هدف کمینه کردن هزینه تطبیق به عنوان یک مساله بهینه‌سازی است [۲]. در روش‌های تخمین آماری مانند فیلتر کالمن^۱ و فیلتر ذره‌ای^۲ از رویکرد فضای حالت^۳ برای مدل کردن ویژگی‌های خاصی از شیء مانند موقعیت، سرعت و شتاب استفاده می‌کنند [۴]. روش ارائه شده در این مقاله در شاخه الگوریتم‌های ردیابی نقاط بر اساس مدل قطعی قرار می‌گیرد.

۲- بررسی کارهای مرتبط

در مرجع [۵] از الگوریتم ترکیبی KLT برای ردیابی هدف متحرک استفاده شده است. این الگوریتم ردیابی نقاط ویژگی، از بخش‌های استخراج و تطبیق ویژگی و تصحیح خطا تشکیل شده است. نوآوری آن شامل تعریف دو آستانه t_h و N است. آستانه t_h به منظور حذف نقاط کم اهمیت مورد استفاده قرار می‌گیرد و با مقدار نصف اندازه ناحیه محصور کننده هدف مقداردهی می‌شود. در صورتی که فاصله اقلیدسی نقاط کلیدی از مرکز هدف، از آستانه t_h بیشتر باشد، آن نقطه کلیدی حذف خواهد شد. همچنین در فرآیند ردیابی هرگاه تعداد نقاط کلیدی از آستانه N کم‌تر باشد، از ناحیه هدف در قاب جاری نقاط کلیدی جدید استخراج می‌شود.

در مرجع [۳] یکی از الگوریتم‌های ردیابی مبتنی بر تطبیق مشخصه ارائه شده است. در قاب اول نقاط کلیدی توسط آشکارساز SURF از ناحیه هدف که به صورت دستی مشخص شده است، استخراج می‌شود. سپس در قاب‌های بعدی نقاط کلیدی توسط آشکارساز SURF پیرامون ناحیه هدف در قاب قبلی و با شعاعی بیشتر استخراج می‌شود. همچنین در این الگوریتم برای یافتن نقاط تطبیقی و ناحیه هدف از فاصله اقلیدسی بین توصیف‌گرهای نقاط کلیدی استخراج شده استفاده می‌شود.

در مرجع [۶] به منظور ردیابی اهداف متحرک هوایی، با تخمین چگالی هسته گوسی، هیستوگرام وزن دار شده مدل هدف به دست آمده و با اعمال واریانس یک نویز تصادفی در محل هدف، موقعیت ذرات نامزد در قاب بعد پیشگویی می‌شوند. در این مقاله ذرات نامزد توسط فاصله باچاتاریا وزن دار شده و تعداد مراحل باز نمونه برداری در هر قاب، متناسب با وزن ذرات در الگوریتم فیلتر ذره به صورت وقتی تعیین می‌گردند. همچنین به منظور کنترل تغییرات شیء متحرک، شعاع هسته گوسی با لبه‌یابی مشخص می‌شود.

یکی از آشکارسازهای نقاط کلیدی الگوریتم SIFT [۷] است.

4- Affine Scale Invariant Feature Transform (ASIFT)
5- Epipolar
6- Outlier
7- Optimized Random Sampling Algorithm (ORSA)

1- Kalman Filter
2- Particle Filter
3- State Space Models

کلیدی بین تصاویر است که در کاربردهایی نظیر تشخیص شیء، تناظریابی تصاویر، بازیابی شیء در بانک‌های اطلاعاتی چندرسانه‌ای، ربات‌های خودکار محلی، تعیین خسارت وسائط متحرک مورد استفاده قرار می‌گیرد. در حالت آرمانی نقاط کلیدی باید، مشخص و متمایز باشند، به آسانی استخراج شوند و مستقل باشند. به عبارت دیگر نباید وابستگی به مقیاس، دوران، تغییرات در دید سه‌بعدی، تغییرات در روشنایی تصویر و میزان نویز تصویر داشته باشند.

ایده اساسی در الگوریتم ASIFT شبیه‌سازی اعوجاج‌های حاصل از تغییر جهت محور نوری دوربین نسبت به تغییر منظر اخذ تصویر است. در شکل (۲) صفحه تصویر از یک منظر نشان داده شده است. بر اساس هندسه مورد استفاده در الگوریتم ASIFT، زوایای θ و ϕ ، بیانگر امتداد محور نوری دوربین و توصیف‌کننده منظر اخذ تصویر می‌باشند. θ و ϕ به ترتیب به عنوان زوایای طولی^۶ و عرضی^۷ در نظر گرفته می‌شوند. زاویه ψ به عنوان زاویه چرخش محور نوری دوربین حول خود و λ نیز به‌عنوان پارامتر مقیاس در نظر گرفته می‌شود [۹].

در الگوریتم ASIFT جهت شبیه‌سازی تصاویر از منظرهای مختلف، زاویه ϕ متناسب با زاویه θ و با گام‌های $\Delta\phi = 72^\circ/t$ تغییر داده می‌شود. t ، زاویه تیلت^۸ بوده و برابر $t = 1/\cos\theta$ است. زاویه تیلت میزان اعوجاج‌های حاصل از تغییر جهت محور نوری دوربین را مشخص می‌کند. قابل ذکر است که محدوده این زوایا به ترتیب $\phi \in [0, \pi)$ و $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ است [۹]. بنابراین، مطابق با شکل (۲)، با در نظر گرفتن حالت‌های مشخص و محدودی از زوایای دورانی، تصاویر یک شیء از منظرهای مختلف شبیه‌سازی می‌شود.

الگوریتم ASIFT با شبیه‌سازی محور نوری دوربین، تصاویری را با زوایای تیلت مختلف برای استخراج نقاط کلیدی ایجاد می‌کند، که نمونه‌ای از این شبیه‌سازی در شکل (۳) نشان داده شده است. پس از این پردازش به منظور استخراج نقاط کلیدی الگوریتم استاندارد SIFT روی جفت تصاویر شبیه‌سازی شده اعمال می‌شود. این عمل برای دو قاب متوالی انجام شده و در نهایت فرآیند تناظریابی میان نقاط کلیدی توسط الگوریتم ORSA انجام می‌گیرد. در الگوریتم ASIFT برای به‌دست آوردن نقاط تطبیقی توسط الگوریتم ORSA تنها از نقاط کلیدی استخراج شده الگوریتم SIFT استفاده می‌شود و از توصیف‌گرهای SIFT استفاده نمی‌شود. در ادامه نحوه آشکارسازی نقاط کلیدی

یکی از چالش‌های مطرح در الگوریتم‌های ردیابی مقابله با پوشیدگی هدف است. بحث مقابله با پوشیدگی نسبی و آشکارسازی مجدد هدف در الگوریتم‌های ردیابی مبتنی بر هسته بیشتر مطرح است؛ زیرا در این الگوریتم‌ها برای یافتن موقعیت جدید هدف، دنباله‌ای از الگوهای هدف از قاب‌های قبلی ذخیره می‌شود. این امر تغییرات تدریجی هدف را کنترل کرده و در صورت پوشیدگی بخشی از هدف قابلیت یافتن بهترین کاندید را فراهم می‌کند. در مرجع [۱۱] از بازنمایی پراکنده^۱ الگوهای استخراج شده از ناحیه هدف و پس‌زمینه برای ردیابی هدف متحرک استفاده شده است. هدف این الگوریتم یافتن بهترین کاندید با حداقل کردن خطای بازسازی با توجه به الگوهای ذخیره شده از شیء است. در این روش تعداد الگوهایی که از ناحیه هدف و پس‌زمینه استخراج می‌شود بین ۴۰۰ تا ۶۰۰ عدد بوده و قادر است پوشیدگی‌های سنگین را کنترل کند. در مقاله ارائه شده به منظور کنترل پوشیدگی و آشکارسازی مجدد هدف ابتدا ناحیه‌های کاندید حضور هدف با استفاده از روش شناسایی اشیاء برجسته^۲ در تصویر توسط الگوریتم باقیمانده طیفی^۳ استخراج می‌شود، سپس با استفاده از اطلاعات الگوهای استخراج شده از هدف و روش همبستگی متقاطع نرمال شده^۴ موقعیت جدید هدف به‌دست می‌آید.

در بیشتر الگوریتم‌های ردیابی در صورت تغییر مقیاس ناگهانی هدف در اثر دور یا نزدیک شدن شیء نسبت به دوربین، ردیاب با مشکل مواجه می‌شود، برای حل این مشکل در این مقاله از اطلاعات تاریخچه مقیاس هدف در قاب‌های قبلی استفاده خواهد شد.

این مقاله به‌صورت زیر سازماندهی شده است: در بخش سوم الگوریتم ASIFT ارائه می‌شود. در بخش چهارم الگوریتم RANSAC معرفی می‌شود. در بخش پنجم الگوریتم آشکارسازی اشیاء برجسته در تصویر به منظور کنترل پوشیدگی ارائه می‌شود. سپس در بخش ششم روش پیشنهادی بیان می‌شود و روش مورد استفاده در ردیابی نقاط ویژگی توضیح داده خواهد شد. نتایج حاصل از شبیه‌سازی روش پیشنهادی بر روی چند دنباله ویدئویی در بخش هفتم ارائه خواهد شد و در بخش هشتم نتیجه‌گیری بحث می‌شود.

۳- الگوریتم ASIFT

الگوریتم ASIFT روشی برای آشکارسازی و تناظریابی^۵ نقاط

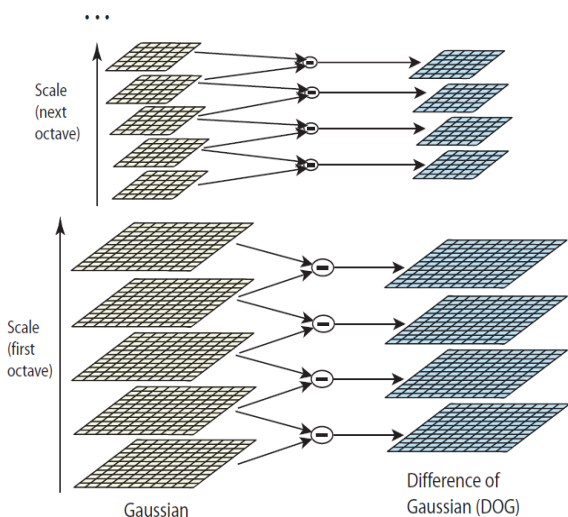
6- Longitude Angle
7- Latitude Angle
8 -Tilt Angle

1- Sparse Representation
2- Saliency Detection
3 -Spectral Residual
4 -Normalized Cross Correlation
5- Matching

بیشینه تصاویر در مقیاس‌های مختلف ایجاد شده و در این مرحله نقاط کلیدی مستقل از مقیاس می‌گردند. در رابطه (۱) نحوه آشکارسازی نقاط کمینه و بیشینه در فضای مقیاس آورده شده است، که در آن، I تصویر اصلی و G تابع هسته گاوسی است که از طریق کانولوشن، تصویر نرم شده L را تولید می‌کند. سپس با کم کردن دو تصویر حاصل شده L در هر دسته پنج‌تایی با تصویر مجاورش، تصویر D بر اساس اختلاف تفاضل گاوسی‌ها ایجاد می‌گردد. به تابع $L(x, y, \sigma)$ فضای مقیاس یک تصویر نیز می‌گویند که از طریق کانولوشن تصویر با یک فیلتر گاوسی به دست می‌آید و در رابطه (۲) نشان داده شده است.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)$$

$$L(x, y, k\sigma) = G(x, y, \sigma) * I(x, y) \quad (2)$$



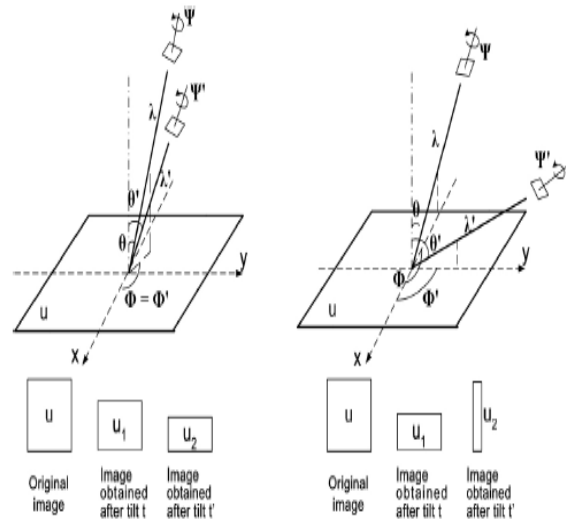
شکل (۴). هرم تصویری با تابع گاوسی [۱۲]

۲-۳- انتخاب نقاط کلیدی از بین نقاط کمینه و

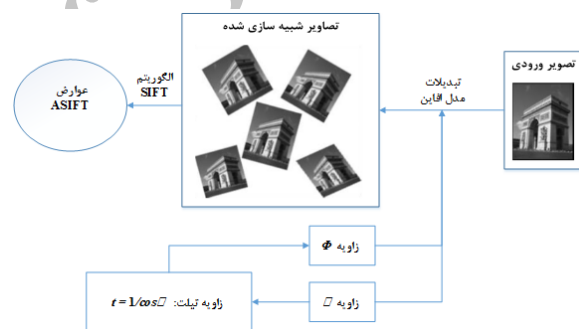
بیشینه اصلی

در این مرحله مقادیر شدت روشنایی سطوح خاکستری هر پیکسل با ۸ پیکسل مجاور آن و با ۹ پیکسل در تصاویر مجاور بالایی و پایینی (که از لحاظ σ با هم اختلاف داشتند و از تابع اختلاف تفاضل گاوسی حاصل شده بود) مقایسه می‌گردد، اگر مقدار این پیکسل از تمام ۲۶ پیکسل همسایه بیش‌تر یا کم‌تر باشد، به عنوان نامزد نقطه کلیدی انتخاب می‌شود، که در شکل (۵) نشان داده شده است. بدیهی است هزینه محاسباتی این مقایسه کم است چون بیش‌تر نقاط در بررسی‌های اولیه حذف می‌شوند.

توسط الگوریتم SIFT آورده شده است.



شکل (۲). تصاویر شبیه‌سازی شده توسط الگوریتم ASIFT [۹]



شکل (۳). الگوریتم استخراج عوارض در الگوریتم ASIFT

۳-۱- آشکارسازی نقاط کمینه و بیشینه در فضای

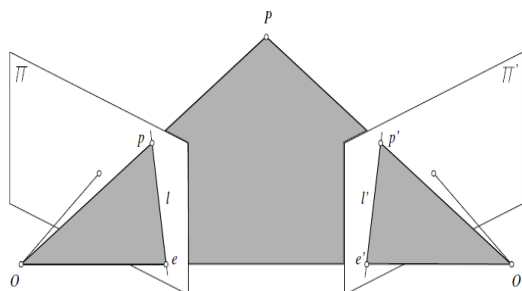
مقیاس ۱

با استفاده از مفاهیم هرم تصویری^۲ و فضای مقیاس، آشکارسازی نقاط کمینه و بیشینه در فضای مقیاس انجام می‌پذیرد. با توجه به مفهوم فضای مقیاس که در آن تصویر در مقیاس‌های مختلف ساخته می‌شود، مطابق شکل (۴) چند اکتاو مختلف تشکیل می‌شود، سپس برای هر اکتاو، پنج تصویر با انحراف مختلف به وجود می‌آید. در هر اکتاو اندازه تصویر نصف می‌شود و دسته‌های پنج‌تایی از طریق ضرب هسته گاوسی در تصویر اصلی با انحراف معیارهای متفاوت تشکیل می‌شود. سپس همان‌گونه که در شکل (۴) ملاحظه می‌گردد، هر دو تصویر مجاور در داخل دسته‌های پنج‌تایی بر اساس جبر ماتریسی از هم کم می‌شوند و تصاویر جدید بر اساس فیلتر اختلاف تفاضلی ایجاد می‌گردد. این مشتق‌گیری ثنوبه در واقع جهت پیدا کردن نقاط کمینه و بیشینه در فضای مقیاس انجام می‌شود. در نتیجه مقادیر کمینه و

1- Scale Space

2- Image Pyramid

نقطه زمینی P در دو تصویر، نقاط p و p' است. این پنج نقطه مربوط به صفحه اپی پولار تولیدشده توسط پرتوهای نوری متقاطع OP و $O'P$ می‌باشند. خطی که مرکز دو دوربین را به هم وصل می‌کند و دو صفحه π و π' را قطع می‌کند خط پایه نامیده می‌شود. خطوط l و l' به ترتیب متناظر با نقاط p و p' بوده و خطوط اپی پولار نامیده می‌شوند و نقطه e تصویر مرکز دوربین اول (O) در صفحه دوربین دوم بوده و نقطه اپی پول نامیده می‌شود. اگر نقاط p و p' تصاویر یک نقطه مشابه در فضای سه بعدی باشند، نقطه p' باید روی خط اپی پولار متناظر با p قرار بگیرد [۱۴].



شکل (۶). قید هندسی اپی پولار [14]

در این قسمت فرض می‌شود پارامترهای درونی دوربین مشخص است، بنابراین، $p = p'$ است. با توجه به قید هندسی اپی پولار سه بردار \vec{Op} ، $\vec{O'p'}$ و $\vec{OO'}$ هم صفحه می‌باشند. حاصل ضرب خارجی دو بردار $\vec{OO'}$ و $\vec{O'p'}$ بردار نرمال صفحه است که بر بردار \vec{Op} عمود است و لذا رابطه (۳) برقرار است.

$$\vec{Op} \cdot |\vec{OO'} \times \vec{O'p'}| = 0 \quad (3)$$

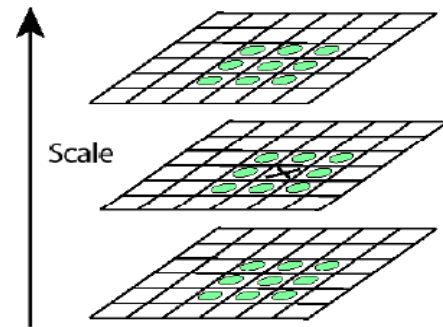
حال رابطه مستقل از مختصات (۳) در مختصات صفحه دوربین اول به صورت رابطه (۴) نوشته می‌شود.

$$p \cdot [t \times (Rp')] = 0 \quad (4)$$

که در آن، $p = (u, v, 1)^T$ و $p' = (u', v', 1)^T$ نشان‌دهنده بردارهای مختصات همگن تصویر بوده، t بردار انتقال $\vec{OO'}$ است که سیستم مختصات دو تصویر را از هم جدا کرده و R ماتریس دوران است. در نهایت رابطه (۴) به صورت رابطه (۵) بازنویسی می‌شود.

$$p^T F p' = 0 \quad (5)$$

در رابطه (۵)، Fp' نشان‌دهنده بردار مختصات خط اپی پولار مربوط به نقطه p' در صفحه تصویر اول، $F^T p$ نشان‌دهنده خط اپی پولار مربوط به نقطه p در صفحه تصویر دوم و F ماتریس بنیادی است. هندسه اپی پولار با استفاده از یک ماتریس 3×3 به عنوان ماتریس بنیادی، F قابل توصیف است. ماتریس F مطابقت میان یک نقطه و خط اپی پولار مربوط به آن را توصیف کرده و



شکل (۵). کمینه و بیشینه محلی حاصل از DOG [۱۲]

در مرحله بعد برای تعیین دقیق موقعیت هر نقطه کلیدی از درون یابی نقاط مجاور استفاده می‌شود و نقاطی که حساس به نویز باشند، مانند نقاط با کنتراست کم و نقاط روی لبه حذف می‌شوند.

۴- الگوریتم RANSAC

الگوریتم RANSAC روشی جهت برازش یک مدل هندسی است که ممکن است تعدادی خط در میان نقاط مربوط به مدل وجود داشته باشد [۱۳]. در این روش یک مجموعه از ماتریس‌های بنیادی به دست آمده از مجموعه‌های تصادفی شامل تطابق‌های ۸ تایی برای یافتن بزرگ‌ترین مجموعه مجتمع استفاده می‌شود. در واقع، ماتریس بنیادی تخمین زده شده از روی مجموعه نقاط تصادفی باید قادر باشد بیشترین تطابق بین نقاط دو مجموعه را فراهم کند. الگوریتم RANSAC دارای دو مشکل، تعداد تکرارهای مورد نیاز برای یافتن بهترین ماتریس بنیادی از روی مجموعه‌های تصادفی شامل تطابق‌های ۸ تایی و آستانه از پیش تنظیم شده برای تطابق نقاط کلیدی است. در روش ORSA با ترکیب اندازه مجموعه‌ای که برای محاسبه ماتریس بنیادی استفاده می‌شود و یافتن دقت مورد نیاز برای تطابق، این مشکل حل می‌شود. الگوریتم ORSA یک توسعه از الگوریتم RANSAC است که برای حذف تناظرهای اشتباه در الگوریتم ASIFT استفاده می‌شود. الگوریتم ORSA قادر به شناسایی حرکات صلب بوده و در حالتی که بیش از ۹۰ درصد داده‌ها را، داده پرت تشکیل دهد، قادر است ماتریس بنیادی را تخمین بزند [۱۰]. در این مقاله از ترکیب دو روش ORSA و RANSAC برای یافتن نقاط تطبیقی و حذف تناظرهای اشتباه استفاده شده است.

۴-۱- هندسه اپی پولار

هندسه اپی پولار، یک قید اساسی برای تصاویر اخذ شده از منظرهای مختلف است. مطابق شکل (۶) دو صفحه تصویری برای دو دوربین با مراکز تصویری O' و O نشان داده شده‌اند. تصویر

در صورتی که ماکزیمم مقدار این فواصل از آستانه، T_E بیش‌تر باشد آن‌گاه جفت تناظر (p, p') خطا بوده و حذف می‌شوند.

۵- آشکارسازی اشیاء برجسته با استفاده از

باقیمانده طیفی

هدف از آشکارسازی، یافتن اشیاء برجسته و استخراج آن‌ها از پس‌زمینه است. در مقاله ارائه شده برای آشکارسازی و مقابله با پوشیدگی نسبی از روش شناسایی اشیاء برجسته ارائه شده در مرجع [۱۵] استفاده می‌شود.

ابتدا Barlow در مرجع [۱۶] نظریه حذف اطلاعات تکراری را پیشنهاد داد. بر اساس این نظریه اطلاعات تصویر بر اساس رابطه (۹) از دو بخش تشکیل شده است.

$$H(\text{Image}) = H(\text{Innovation}) + H(\text{Prior Knowledge}) \quad (9)$$

که در آن، $H(\text{Innovation})$ نشان‌دهنده اطلاعات جدید و $H(\text{Prior Knowledge})$ اطلاعات تکراری است که باید توسط سامانه حذف شود. این اطلاعات تکراری مربوط به ویژگی‌های تغییرناپذیر آماری از محیط است. همان‌طور که در مرجع [۱۵] مشخص شده، با حذف اطلاعات آماری اجزای تکراری در تصویر، اشیاء جدید استخراج می‌شوند.

اگر تصویر به عنوان یک سیگنال دو بعدی در نظر گرفته شود، تبدیل فوریه تصویر از دو بخش حقیقی و موهومی تشکیل شده است؛ لذا برای تصویر $I(x)$ روابط (۱۰) برقرار است.

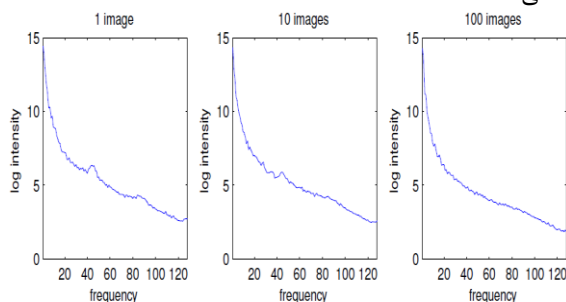
$$\mathcal{A}(f) = \Re(\mathcal{F}[I(x)])$$

$$\mathcal{P}(f) = L(\mathcal{F}[I(x)]) \quad (10)$$

و لگاریتم دامنه فرکانس توسط رابطه (۱۱) محاسبه می‌شود.

$$L(f) = \log(\mathcal{A}(f)) \quad (11)$$

همان‌طور که در شکل (۷) نشان داده شده است، لگاریتم دامنه فرکانس تصاویر مختلف در حوزه فوریه دارای رفتار مشابه و خطی است.



شکل (۷). میانگین طیفی برای ۱۰۰، ۱۰ و ۱ تصویر [۱۵]

به‌منظور تخمین شکل کلی نمودار میانگین طیفی تصاویر، از

میان این ماتریس و موقعیت نقاط کلیدی متناظر در دو تصویر، رابطه (۵) برقرار است [۱۴].

باید توجه شود که قید هندسی اپی‌پولار آورده شده در رابطه (۵) یک معادله خطی با ۹ ضریب برای ماتریس بنیادی است که به صورت رابطه (۶) بیان می‌شود.

$$(x, y, 1) \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = 0$$

$$\Leftrightarrow (xx', xy', x, yx', yy', y, y', yx', 1) \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0 \quad (6)$$

در تبدیلات همگن عنصر $F_{33} = 1$ در نظر گرفته می‌شود. حال با استفاده از ۸ نقطه متناظر $p_i \leftrightarrow p'_i$ ($i = 1, 2, \dots, 8$) کمینه‌سازی تخمین حداقل مربعات خطی ارائه شده در رابطه (۷) ماتریس F تخمین زده می‌شود [۱۴].

$$\sum_{i=1}^n (p_i^T F p'_i)^2 \quad (7)$$

با فرض این که مدل هندسی مورد نظر دارای u پارامتر مجهول بوده و تعداد حداقل m متناظر از میان n متناظر موجود برای $(n > m)$ برای برآورد آن لازم باشد که در این‌جا ۸ متناظر برای برآورد ماتریس بنیادی مورد نیاز است، آن‌گاه در الگوریتم RANSAC به ترتیب زیر برای برآورد پارامترهای مجهول عمل می‌شود.

۱- تعداد $m = 8$ جفت از میان n جفت موجود به صورت تصادفی انتخاب می‌شود.

۲- با استفاده از $m = 8$ جفت تناظرهای انتخاب شده، پارامترهای مجهول مدل هندسی برآورد می‌شود.

۳- سازگاری $n - m$ جفت تناظر باقیمانده با مدل هندسی برآورد شده، بررسی می‌شود و تعداد جفت تناظرهایی که با آن سازگار هستند، با در نظر گرفتن یک حد آستانه تعیین می‌شوند $SSD(p'_i, Fp_i) < \epsilon$.

۴- مراحل ۱ تا ۳، k مرتبه تکرار شده و در نهایت راه حلی که بیشترین تعداد جفت تناظرهای سازگار با مدل هندسی را حاصل کند، به عنوان جواب انتخاب می‌شود.

بعد از برآورد ماتریس F با استفاده از الگوریتم RANSAC، برای هر جفت تناظر اولیه نظیر (p, p') فاصله آن‌ها از خط اپی‌پولار مربوطه با استفاده از رابطه (۸) محاسبه می‌شود.

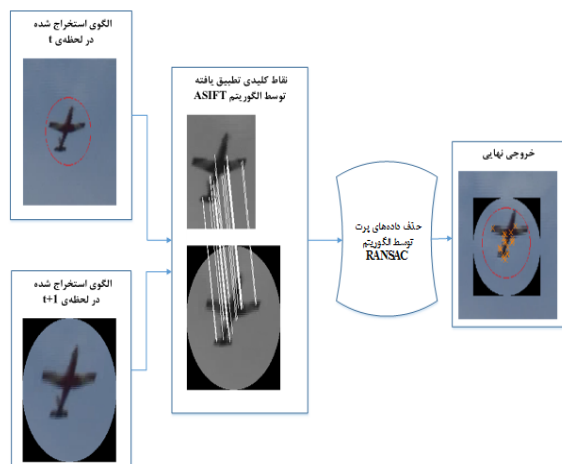
$$d_q = \frac{|p^T F p'|}{\sqrt{(Fp)_1^2 + (Fp)_2^2}} \quad (8)$$

۶- روش پیشنهادی

در الگوریتم پیشنهادی ابتدا در قاب اول (در لحظه t) ناحیه هدف به صورت دستی مشخص شده و دایره‌ای پیرامون هدف رسم می‌شود. در قاب بعدی (لحظه $t + 1$) با مرکزیت ناحیه دایره‌ای مشخص شده در قاب اول و با شعاعی بزرگ‌تر، یک الگو استخراج می‌شود. سپس با استفاده از الگوریتم ASIFT نقاط کلیدی استخراج شده و تطبیق‌های بین این دو قاب با استفاده از الگوریتم ORSA محاسبه می‌شود. با توجه به این‌که تعدادی از تناظرهای به دست آمده اشتباه می‌باشند، با استفاده از الگوریتم RANSAC این تناظرهای اشتباه تعیین شده و حذف می‌شوند. سپس پیرامون نقاط تطبیقی به دست آمده دایره‌ای رسم می‌شود که نشان دهنده‌ی موقعیت هدف در قاب جاری خواهد بود. در شکل (۹) نمای کلی الگوریتم ردیابی ارائه شده نشان داده شده است.

معیار ارزیابی جهت تطبیق نقاط کلیدی، آستانه $T_{KeyPoints}$ است که با توجه به تغییر مقیاس هدف در قاب‌های متوالی مشخص می‌شود. اگر پس از افزایش شعاع ناحیه جستجو تطبیق مناسبی برای هدف فراهم نشود، الگوریتم در فاز آشکارسازی قرار می‌گیرد. در نهایت ناحیه مشخص شده برای هدف در قاب جاری به عنوان الگوی معیار برای تعیین ناحیه هدف در قاب بعدی ذخیره می‌شود و مراحل ذکر شده تکرار خواهد شد.

به منظور کنترل پوشیدگی نسبی هدف و کاهش خطای ردیابی در صورتی که تطبیق مناسبی برای هدف فراهم نشود، با استفاده از روش آشکارسازی اشیاء برجسته، آشکارسازی مجدد انجام می‌شود. در این حالت اگر هدف پنهان شود یا پوشیدگی نسبی رخ دهد، به روز رسانی الگوی هدف انجام نمی‌شود و تا زمانی که هدف مجدد در دید دوربین قرار بگیرد و تطبیقی برای آن یافت شود، الگوی پیشین هدف حفظ خواهد شد.



شکل (۹). نمای کلی الگوریتم ردیابی

کانوالو یک فیلتر میانگین $h_n(f)$ در لگاریتم دامنه فرکانس $L(f)$ استفاده می‌شود که در رابطه (۱۲) مشخص شده است.

$$\mathcal{A}(f) = h_n(f) * L(f) \quad (12)$$

در مرجع [۱۵]، $h_n(f)$ یک ماتریس $n \times n$ به صورت رابطه (۱۳) تعریف شده است.

$$h_n(f) = \frac{1}{n^2} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \quad (13)$$

حال با حذف اطلاعات آماری تکراری از تصویر، اشیاء برجسته نمایان می‌شوند که در رابطه (۱۴) مشخص شده است.

$$\mathcal{R}(f) = L(f) - \mathcal{A}(f) \quad (14)$$

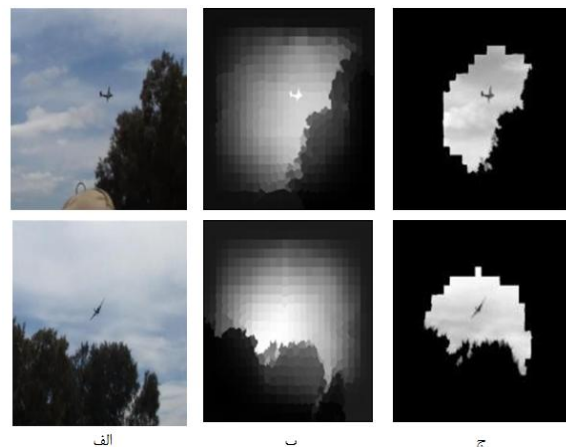
در رابطه (۱۴)، $\mathcal{R}(f)$ باقیمانده طیفی نامیده می‌شود. پس از تشکیل نقشه برجستگی از روی باقیمانده طیفی، برای حذف خطاهای احتمالی از ضرب یک فیلتر گاوسی در معکوس تبدیل فوریه استفاده می‌شود که در رابطه (۱۵) مشخص شده است.

$$S(x) = g(x) * \mathcal{F}^{-1}[\exp(\mathcal{R}(f) + \mathcal{P}(f))]^2 \quad (15)$$

در نهایت به منظور تشکیل ماسک خروجی از اشیاء موجود در تصویر، عملیات آستانه‌گذاری روی نقشه‌ی برجستگی به صورت رابطه (۱۶) انجام می‌شود.

$$O(x) = \begin{cases} 1 & \text{if } S(x) > \text{threshold} \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

و مقدار آستانه به صورت $\text{threshold} = E(S(x)) \times 3$ نظر گرفته می‌شود که $E(S(x))$ میانگین شدت روشنایی نقشه برجستگی است. در شکل (۸) خروجی حاصل از اعمال الگوریتم آشکارسازی اشیاء برجسته آورده شده است.



شکل (۸). استخراج نقشه برجستگی با استفاده از باقیمانده طیفی، (الف) تصاویر اصلی، (ب) باقیمانده طیفی، (ج) ناحیه کاندید تصاویر آستانه‌گذاری شده

۶-۱- استفاده از اطلاعات مقیاس سوژه

جدول (۱). میزان هم پوشانی ناحیه محصور کننده هدف

نام دنباله ویدئویی	تعداد الگوهای ذخیره شده از شعاع هدف			
	۵	۱۰	۱۵	۲۰
aerobatics_2	۰/۸۱	۰/۹۳	۰/۸۸	۰/۷۹
multiple_2	۰/۸۳	۰/۹۴	۰/۸۹	۰/۷۶
big_2	۰/۷۹	۰/۹۰	۰/۸۵	۰/۷۱

پس از استخراج الگوی جدید در قاب جاری و یافتن نقاط تطبیقی توسط الگوریتم ASIFT و همچنین حذف تناظرهای اشتباه توسط الگوریتم RANSAC، مرکز ناحیه محصور کننده هدف از روی نقاط تطبیق یافته به صورت رابطه (۱۷) محاسبه می شود.

$$Center_c = \frac{1}{N} \sum_{j=1}^N KeyPoint_j \quad (17)$$

در رابطه (۱۷) $Center_c$ تخمینی از مختصات مرکز شی در قاب جاری، $KeyPoint_j$ مختصات نقطه کلیدی زام در قاب جاری و N تعداد کل نقاط کلیدی تطبیق یافته در قاب جاری را مشخص می کند. پس از محاسبه مرکز شیء طبق رابطه (۱۷)، بیشترین فاصله اقلیدسی بین نقاط تطبیق یافته در قاب جاری با $Center_c$ به عنوان شعاع دایره محصور کننده در نظر گرفته می شود. رابطه (۱۸) نحوه محاسبه شعاع دایره محصور کننده از روی نقاط کلیدی تطبیق یافته را نشان می دهد.

$$R_m = \max \|Center_c - KeyPoint_j\|_2, j = 1, 2, \dots, N \quad (18)$$

در رابطه (۱۸)، $\|Center_c - KeyPoint_j\|_2$ مشخص کننده فاصله اقلیدسی بین مرکز دایره و نقطه کلیدی زام در قاب جاری است.

پس از محاسبه شعاع ناحیه محصور کننده هدف R_m از روی نقاط کلیدی تطبیق یافته در قاب جاری، فاصله اقلیدسی این شعاع با هر یک از شعاع های $R_1, R_2, R_3, \dots, R_{10}$ از قاب های قبلی به منظور به دست آوردن نزدیک ترین شعاع، طبق رابطه (۱۹) محاسبه می شود.

$$R_{similar} = \arg \min_i |R_m - R_i|, i = 1, 2, \dots, 10 \quad (19)$$

حال $R_{similar}$ در ابتدای دنباله R_1, R_2, \dots, R_{10} قرار می گیرد تا اثر بخشی بیشتری در تعیین شعاع ناحیه محصور کننده هدف داشته باشد. همچنین فاصله اقلیدسی R_m با هر یک از شعاع های $R_1, R_2, R_3, \dots, R_{10}$ به منظور به دست آوردن شعاعی با بیشترین اختلاف، طبق رابطه (۲۰) محاسبه می شود.

$$R_{dissimilar} = \arg \max_i |R_m - R_i|, i = 1, 2, \dots, 10 \quad (20)$$

سپس مقدار R_m با $R_{Dissimilar}$ جایگزین می شود. در نهایت شعاع ناحیه محصور کننده هدف در قاب جاری طبق رابطه (۲۱)

هر اندازه میزان انطباق ناحیه محصور کننده هدف بر آن بیش تر باشد خطای بازسازی ناحیه هدف کم تر خواهد شد. یکی از مشکلاتی که در اکثر الگوریتم های ردیابی وجود دارد این است که از اطلاعات تاریخیچه مقیاس هدف در قاب های قبلی استفاده نمی شود. به همین دلیل با تغییر مقیاس هدف در اثر دور یا نزدیک شدن هدف نسبت به دوربین، ناحیه محصور کننده هدف که به صورت چند ضلعی یا دایره ای است، با خطای زیادی مواجه خواهد شد. در این مقاله به جای استفاده از چند ضلعی برای نمایش ناحیه هدف از دایره ای که ناحیه هدف را محصور خواهد کرد استفاده می شود؛ لذا به جای نگهداری طول و عرض چند ضلعی محصور کننده هدف کافی است شعاع دایره محصور کننده هدف به عنوان پیشینه ردیابی نگه داری شود.

برای حل مشکل مطرح شده، در این مقاله از اطلاعات تاریخیچه مقیاس هدف در قاب های قبلی استفاده می شود. برای این منظور در قاب های متوالی ۱۰ الگو از شعاع دایره هایی که هدف را محصور می کنند به صورت $R_1, R_2, R_3, \dots, R_{10}$ ذخیره می شود. در این الگوها R_i نشان دهنده شعاع ناحیه محصور کننده هدف است. همچنین با توجه به نزدیکی شعاع ناحیه محصور کننده هدف در قاب های متوالی نسبت به مقیاس شعاع هدف در قاب جاری، الگوها مرتب می شوند و برای هر کدام از الگوهای R_i یک ضریب اثر بخشی C_i تعریف می شود. مقدار این ضریب اثر بخشی در بازه $(0-1)$ است. ضریب ۱ متناظر با الگوی R_1 بوده و ضریب $0/1$ متناظر با الگوی R_{10} است. مقادیر ضرایب اثر بخشی در طول فرآیند ردیابی مقادیر ثابت $C = [1, 0.9, 0.8, \dots, 0.1]$ است و تنها الگوهای R_i در هر قاب به روزرسانی می شوند.

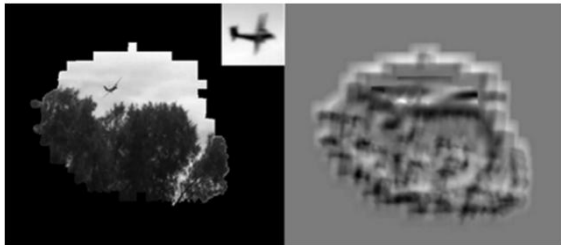
طی آزمایش های انجام شده، اگر تعداد الگوهای ذخیره شده از شعاع هدف در قاب های متوالی کم تر از ۱۰ الگو باشد، تخمین خوبی از تغییر مقیاس هدف به دست نمی آید، علاوه بر این افزایش تعداد الگوهای ذخیره شده منجر به افزایش خطا خواهد شد. در جدول (۱) برای نشان دادن تاثیر تعداد الگوهای ذخیره شده از شعاع هدف در میزان هم پوشانی ناحیه محصور کننده هدف با خروجی حقیقی، الگوریتم پیشنهادی روی سه دنباله ویدئویی اعمال شد. همان طور که در جدول (۱) مشخص شده است، در صورتی که ۱۰ الگو از شعاع هدف در قاب های متوالی برای تخمین مقیاس هدف در قاب جاری استفاده شود، میزان هم پوشانی ناحیه محصور کننده هدف با خروجی حقیقی بیشتر خواهد شد.

که در رابطه (۲۲)، f تصویر اصلی، \bar{t} میانگین الگوی مورد مقایسه و $\bar{f}_{u,v}$ نشان دهنده میانگین تصویر اصلی درون ناحیه‌ای از الگوی t است که در جهت (u, v) جابه‌جا شده است و با استفاده از رابطه (۲۳) محاسبه می‌شود.

$$\bar{f}_{u,v} = \frac{1}{N_x N_y} \sum_{x=u}^{x+N_x-1} \sum_{y=v}^{y+N_y-1} f(x, y) \quad (23)$$

در رابطه (۲۳)، γ نقشه همبستگی متقاطع نرمال شده را مشخص می‌کند، که در شکل (۱۱) قسمت ب نشان داده شده است.

همان‌طور که در شکل (۱۱) نشان داده شده است، با وجود تغییر ظاهر هدف، الگوی میانگین ذخیره شده از پیشینه هدف در نقشه اطمینان بیشترین بازنمایی را دارد. همچنین مشخص شده است، در صورتی که پس‌زمینه هدف پیچیده باشد قابلیت آشکارسازی مجدد وجود دارد.



شکل (۱۱). آشکارسازی، با استفاده از پیشینه هدف

الگوریتم ردیابی ارائه شده به شرح زیر است:

- ۱- مشخص کردن موقعیت هدف در قاب اول (در لحظه t) به صورت دستی با استفاده از موس و تعیین الگوی معیار
- ۲- استخراج الگوی جدید در لحظه $t + 1$ با مرکزیت ناحیه دایره‌ای مشخص شده در قاب قبلی و با شعاعی بزرگ‌تر
- ۳- استخراج نقاط کلیدی و یافتن نقاط تطبیق‌یافته بین الگوها در لحظه‌های t و $t + 1$ ، با استفاده از الگوریتم ASIFT
- ۴- حذف تطبیق‌های نادرست با استفاده از الگوریتم RANSAC
- ۵- اگر تعداد تطبیق‌های به‌دست‌آمده از آستانه $T_{KeyPoints}$ بیشتر بود:
 - محاسبه مرکز دایره محصورکننده هدف از طریق میانگین‌گیری بین موقعیت پیکسلی نقاط تطبیق یافته و جایگزینی با مرکز قبلی طبق رابطه (۱۷)
 - محاسبه فاصله اقلیدسی مرکز به‌دست‌آمده در مرحله قبل با نقاط کلیدی تطبیق یافته و انتخاب بیشترین فاصله به‌عنوان شعاع دایره، به منظور در برگرفتن تمام نقاط کلیدی تطبیق یافته طبق رابطه (۱۸)

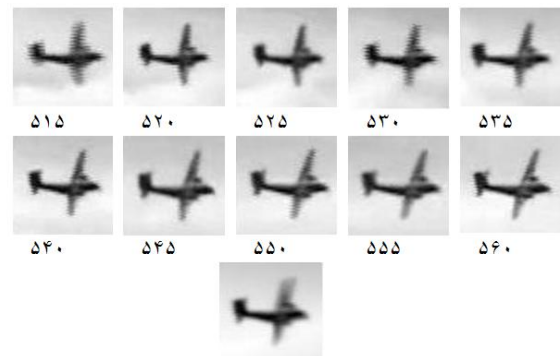
محاسبه می‌شود.

$$R_{current} = \frac{\sum_{i=1}^{ns} C_i R_i}{\sum_{i=1}^{ns} C_i} \quad (21)$$

در رابطه (۲۱)، ns تعداد شعاع‌های قرار گرفته در پشته را مشخص می‌کند.

۶-۲- استفاده از پیشینه هدف

در این مقاله برای کنترل انسداد و تغییرات ایجاد شده در ظاهر هدف، الگوهای استخراج شده از ناحیه هدف توسط الگوریتم ردیابی، نمونه‌برداری می‌شوند. تعداد این الگوها ۱۰ عدد است که در یک پشته نگهداری می‌شوند. همچنین به منظور قابلیت تعمیم و کنترل تغییرات ظاهری، این الگوهای معیار در هر ۵ قاب به‌روزرسانی خواهند شد. نحوه به‌روزرسانی به این گونه است که الگوی هدف در قاب جاری با الگوی ذخیره شده در ابتدای پشته جایگزین می‌شود این جایگزینی به ترتیب ورود خواهد بود. هنگامی که تطبیق مناسبی برای الگوی استخراج شده از ناحیه هدف در قاب جاری و قاب قبلی فراهم نشد، الگوریتم در فاز آشکارسازی قرار می‌گیرد و توسط آشکارسازی اشیاء برجسته با استفاده از باقیمانده طیفی که در بخش پنجم ارائه شد، ناحیه‌های کاندید برای هدف استخراج می‌شود. سپس میانگین ۱۰ الگوی معیار ذخیره شده در پشته محاسبه می‌شود که در شکل (۱۰) نشان داده شده است.



شکل (۱۰). محاسبه تصویر میانگین از روی الگوهای ذخیره شده از هدف

در مرحله بعد احتمال رویت تصویر میانگین محاسبه شده، در ناحیه کاندید به‌دست‌آمده از الگوریتم آشکارسازی اشیاء برجسته با استفاده از باقیمانده طیفی، توسط روش همبستگی متقاطع نرمال شده که در فرمول (۲۲) نشان داده شده است، محاسبه می‌شود.

سپس بهترین تطبیق برای هدف در قاب جاری از روی پیک نقشه، همبستگی متقاطع نرمال شده به‌دست می‌آید.

$$\gamma = \frac{\sum_{x,y} (f(x,y) - \bar{f}_{u,v})(t(x-u,y-v) - \bar{t})}{\sqrt{\sum_{x,y} (f(x,y) - \bar{f}_{u,v})^2 \sum_{x,y} (t(x-u,y-v) - \bar{t})^2}} \quad (22)$$

توجه شود که الگوریتم ASIFT به طور هم‌زمان روی دو تصویر برای استخراج و تطبیق نقاط کلیدی اعمال می‌شود و دیگر احتیاجی به توصیف‌گرهای نقاط کلیدی به‌دست آمده از الگوریتم SIFT نمی‌باشد. نتایج مقایسه زمان اجرای الگوریتم ارائه شده و مقاله‌های [۱، ۱۷ و ۱۸] در جدول (۲) مشخص شده است.

جدول (۲). سرعت اجرا بر حسب فریم بر ثانیه

Video sequence	Proposed	[۱]	[۱۸]	[۱۷]
aerobatics_1	۰/۶۳	۴۱	۹	۱۷/۲
aerobatics_2	۰/۶۷	۳۵	۱۲	۲۰/۴
big_1	۰/۷۱	۴۵/۱	۴/۸	۱۵/۳
big_2	۰/۷۵	۴۰/۸	۵/۱	۱۵
multiple_1	۰/۸۱	۳۸/۲	۳	۱۸/۹
multiple_2	۰/۸۴	۴۷/۴	۳/۹	۱۶/۵
multiple_smoke_1	۰/۶۰	۳۶	۱۳	۱۴/۴
multiple_smoke_2	۰/۶۱	۴۴	۱۴/۸	۱۹/۳
occlusion_1	۰/۹	۳۸	۷	۱۳

در این مقاله از پایگاه داده ارائه شده در [۵] استفاده شده است. این پایگاه داده شامل ۱۴ دنباله ویدئویی است. چالش‌های موجود در این پایگاه داده شامل؛ حرکت دوربین، حرکت سریع بالگردها، پوشیدگی هدف توسط ابرها، درختان و دود منتشر شده از بالگردها، بازتاب نور از سطح شفاف بالگردها و پس‌زمینه پیچیده در پروازهای با ارتفاع پایین، است.

در شکل (۱۲) نتایج اعمال الگوریتم ارائه شده روی ۲ دنباله ویدئویی آورده شده است. در این تصاویر مستطیل مشکی رنگ ناحیه‌ی جستجو را مشخص می‌کند و دایره قرمز رنگ خروجی الگوریتم ردیابی ارائه شده را تعیین می‌کند. هم‌چنین نقاط تطبیق یافته با رنگ سبز مشخص شده‌اند. در پایگاه داده اول، حرکت دوربین، پس‌زمینه پیچیده در ارتفاع پایین، تغییر زاویه دید، دوران و تغییر مقیاس وجود دارد. در پایگاه داده دوم حرکت دوربین، حرکت سریع بالگرد، دوران، تغییر مقیاس و تغییر زاویه دید همراه با تاری وجود دارد.

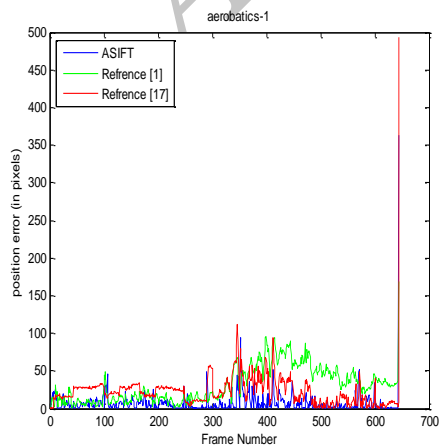
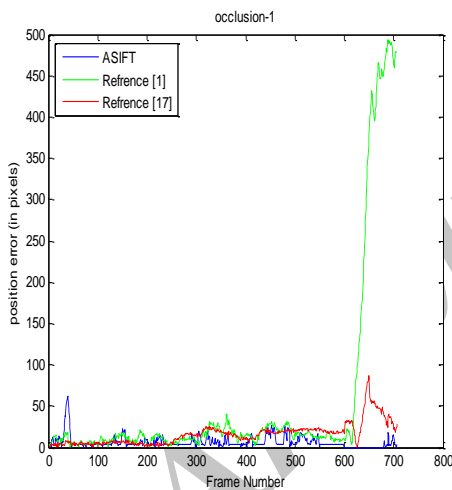
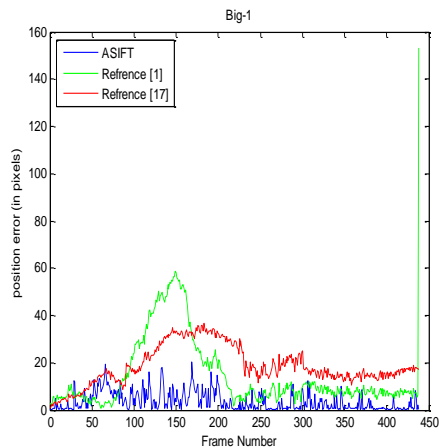
همان‌طور که در شکل (۱۲) مشخص شده است، یکی از مشکلاتی که در روش‌های ردیابی مبتنی بر تطبیق مشخصه وجود دارد این است که نقاط ویژگی استخراج شده از پس‌زمینه، الگوریتم ردیابی را با مشکل مواجه کرده و خطای ردیابی را افزایش می‌دهد. مشکل دیگری که در روش‌های ردیابی مبتنی بر تطبیق مشخصه وجود دارد، مسدود شدن نقاط ویژگی استخراج شده از ناحیه هدف به دلیل پس‌زمینه پیچیده یا پنهان شدن هدف است که الگوریتم ردیابی با مشکل مواجه می‌شود. در این مقاله برای حل این مشکل هنگامی که تعداد تطبیق‌های به دست آمده از ناحیه هدف از آستانه $T_{KeyPoints}$ کم‌تر شد، الگوریتم در فاز آشکارسازی قرار می‌گیرد تا از ایجاد خطا در ردیابی جلوگیری

- در صورتی که شماره قاب کم‌تر از ۱۰ بود، قرار دادن شعاع به‌دست‌آمده از نقاط کلیدی تطبیق یافته R_m در ابتدای پشته
- رسم دایره با استفاده از مرکز به‌دست‌آمده از رابطه (۱۷) و شعاع محاسبه شده طبق رابطه (۲۱) روی قاب جاری، به منظور مشخص کردن ناحیه هدف
- در صورتی که شماره قاب از ۱۰ بیش‌تر بود، محاسبه فاصله اقلیدسی شعاع به‌دست‌آمده از نقاط کلیدی تطبیق یافته R_m با مقادیر شعاع‌های قرار گرفته در پشته $R_1, R_2, R_3, \dots, R_{10}$ از قاب‌های قبلی به منظور به‌دست آوردن شعاع‌های با حداقل و حداکثر اختلاف طبق روابط (۱۹) و (۲۰) و مرتب‌سازی شعاع‌های داخل پشته
- ۶- اگر تعداد تطبیق‌های به‌دست‌آمده کم‌تر از آستانه $T_{KeyPoints}$ بود، چون امکان انسداد نسبی هدف یا پوشیدگی کامل وجود دارد، الگوریتم در فاز آشکارسازی قرار می‌گیرد تا بهترین تطبیق استخراج شود:
- استخراج ناحیه کاندید هدف توسط آشکارسازی اشیاء برجسته با استفاده از باقیمانده طیفی
- محاسبه میانگین ۱۰ الگوی معیار ذخیره شده در پشته
- تخمین موقعیت هدف توسط روش همبستگی متقاطع نرمال شده طبق رابطه (۲۲)
- ۷- خواندن قاب بعدی و ادامه مرحله دوم

۷- آزمایش‌ها

الگوریتم پیشنهادی با استفاده از نرم‌افزار متلب نسخه ۸/۳ و بر روی سامانه‌ای با مشخصات Intel Core i7 با قدرت پردازش ۲/۲ GH و با حافظه اصلی ۶ GB پیاده‌سازی و اجرا شد. نتایج حاصل، قوی بودن روش ارائه شده را در برابر دوران، تغییر مقیاس و تغییر جهت دید نشان می‌دهد. برای ارزیابی زمان اجرای الگوریتم دو قاب متوالی از یک ویدئو با ابعاد 100×100 انتخاب شد و از این تصاویر برای ارزیابی زمان اجرای الگوریتم ارائه شده که شامل الگوریتم ASIFT برای استخراج نقاط کلیدی و یافتن نقاط تطبیقی، الگوریتم RANSAC برای حذف تناظرهای اشتباه، آشکارسازی اشیاء برجسته با استفاده از روش باقیمانده طیفی و اطلاعات پیشینه هدف و استفاده از تاریخچه مقیاس سوژه است، استفاده شد. زمان اجرای روش ارائه شده روی این تصاویر برای استخراج نقاط کلیدی و تطبیق یافته آن‌ها، $3/16$ ثانیه است.

است، آورده شده است. در دنباله‌های ویدئویی aerobatics_specularities_1.aerobatics_clouds_1.aerobatics_3.small_clouds_1 و small_1 به دلیل مات بودن تصویر الگوریتم ارائه شده قادر به استخراج و تطبیق نقاط کلیدی نمی‌باشد. زیرا در این دنباله‌های ویدئویی گرادیان تصویر ضعیف می‌باشد و الگوریتم SIFT نمی‌تواند نقاط کلیدی از ناحیه هدف استخراج کند.



شکل (۱۳). مقایسه میزان خطای پیکسل مرکزی برای سه دنباله ویدئویی با روش‌های [۱] و [۱۷]

به عمل آید.

یکی از معیارهای ارزیابی الگوریتم‌های ردیابی، مشخص کردن میانگین خطای پیکسل مرکزی و میزان همپوشانی ناحیه هدف با آن چه باید در واقعیت رخ دهد است [۱۹]. برای نشان دادن برتری روش ارائه شده در ردیابی اشیاء صلب خطای پیکسل مرکزی برای سه دنباله ویدئویی با روش ارائه شده در مقاله‌های [۱] و [۱۷] مقایسه شد.

خط سیر مرکز هدف، حاصل خروجی الگوریتم ارائه شده از قاب‌های ویدئویی استخراج شد و فاصله اقلیدسی آن با خروجی حقیقی مرکز هدف، به دست آمد که نتایج در شکل (۱۳) مشخص شده است. در دنباله Big-1 بین قاب‌های ۱۰۰ تا ۲۰۰ به دلیل تغییر مقیاس هدف به علت نزدیک شدن به دوربین، خطای ردیابی افزایش می‌یابد؛ ولی در الگوریتم ارائه شده سعی شده است با استفاده از اطلاعات شعاع هدف در قاب‌های قبلی این خطا به حداقل برسد. در دنباله Occlusion-1 بین قاب‌های ۶۰۰ تا ۶۸۰ هدف پشت درختان پنهان می‌شود و الگوریتم ردیابی مرجع [۱] به دلیل ناتوانی در کنترل پوشیدگی، هدف را از دست می‌دهد اما در روش ارائه شده چون تطبیقی برای هدف فراهم نشده است، الگوریتم در فاز آشکارسازی قرار گرفته و خطایی رخ نمی‌دهد؛ همچنین در الگوریتم ردیابی مرجع [۱۷] چون آشکارسازی و ردیابی به صورت هم‌زمان انجام می‌شود قادر است پوشیدگی هدف را کنترل کند. در نهایت در دنباله aerobatics-1 عملکرد الگوریتم ارائه شده نسبت به الگوریتم‌های ردیابی مراجع [۱] و [۱۷] بهتر است.



شکل (۱۲). تصاویر خروجی الگوریتم ردیابی روی دنباله‌های Big-1 و aerobatics-1

در جدول (۳) میانگین خطای پیکسل مرکزی برای ۹ دنباله ویدئویی که خروجی الگوریتم‌های ردیابی مرجع [۱۷] و روش ارائه شده که در شاخه الگوریتم‌های ردیابی نقاط براساس مدل قطعی قرار می‌گیرند و الگوریتم‌های ردیابی مراجع [۱]، [۱۸] که یک ردیاب قوی در میان روش‌های ردیابی براساس هسته

شدن کامل هدف از الگوریتم آشکارسازی اشیاء برجسته و اطلاعات پیشینه هدف استفاده شد و برای کنترل تغییر مقیاس هدف در این مقاله ایده استفاده از تاریخچه مقیاس سوژه ارائه شد.

از جمله مزایای روش ارائه شده می‌توان به کنترل تغییر مقیاس هدف در صورت دور یا نزدیک شدن هدف نسبت به دوربین و کنترل پوشیدگی هدف اشاره کرد. هم‌چنین به دلیل استفاده از الگوریتم ASIFT، در صورت تغییر زاویه دید، انتقال، دوران و تغییر مقیاس هدف نقاط تطبیقی با دقت خوبی به دست می‌آیند. نتایج کیفی و کمی به دست آمده از شبیه‌سازی روش پیشنهادی، نشان می‌دهد در حالی که رزولوشن تصویر مناسب باشد، الگوریتم ارائه شده قادر است هدف را با دقت خوبی ردیابی کند.

۹- مراجع

- [1] K. Zhang, L. Zhang, M. H. Yang, and D. Zhang, "Fast tracking via spatio-temporal context learning," in Computer Vision-ECCV, ed: Springer, pp. 127-141, 2013.
- [2] A. Yilmaz, O. Javed, and M. Shah, "Object Tracking: A Survey," Acm computing surveys, vol. 38, p. 13, 2006.
- [3] H. Shuo, W. Na, and S. Huajun, "Object Tracking Method based on SURF," AASRI Procedia, vol. 3, pp. 351-356, 2012.
- [4] Y. Bar-Shalom, "Tracking and Data Association," Academic Press Professional, Inc, 1987.
- [5] A. S. Mian, "Realtime Visual Tracking of Aircrafts," Techniques and Applications in Digital Image Computing, pp. 351-356, 2008.
- [6] A. Abiri and M. R. Mahzoun, "Aerial Moving Target Tracking using Kernel Density Estimation Based on Particle Filter Algorithm," Tabriz Journal of Electrical Eng., vol. 45, pp. 97-107, 2014 (In Persian).
- [7] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International journal of computer vision, vol. 60, pp. 91-110, 2004.
- [8] J. Morel and G. Yu, "On the Consistency of the SIFT Method," Technical report, ENS Cachan France, 2008.
- [9] J.-M. Morel and G. Yu, "ASIFT: A new framework for Fully Affine Invariant Image Comparison," SIAM Journal on Imaging Sciences, vol. 2, pp. 438-469, 2009.
- [10] X. Li and J. Wang, "Image Matching Techniques for Vision-based indoor navigation systems: a 3D map-based Approach," Journal of Location Based Services, vol. 8, pp. 3-17, 2014.
- [11] X. Jia, H. Lu, and M.-H. Yang, "Visual Tracking via Adaptive Structural Local Sparse Appearance Model," IEEE Conference on Computer vision and pattern recognition, 2012, pp. 1822-1829.
- [12] R. Oji, "An Automatic Algorithm for Object Recognition and Detection based on ASIFT Keypoints," International Journal of Image Processing, vol. 3, p. 29, 2012.
- [13] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Communications of the ACM, vol. 24, pp. 381-395, 1981.
- [14] O. Faugeras, Q. Luong, and T. Papadopoulos, "The Geometry of Multiple Images: The Laws That Govern the Formation of

در صورت تغییر مقیاس هدف، استفاده از اطلاعات مقیاس سوژه باعث می‌شود ناحیه هدف با دقت خوبی استخراج شود. برای نشان دادن برتری روش ارائه شده در کنترل تغییر مقیاس سوژه، میزان هم‌پوشانی ناحیه هدف با خروجی حقیقی، حاصل الگوریتم‌های ردیابی مراجع [۱]، [۱۷] و [۱۸] و روش ارائه شده در جدول (۱) نشان داده شده است.

جدول (۳). میانگین خطای پیکسل مرکزی

Video sequence	ASIFT_T racker	[۱۷]	[۱]	[۱۸]
aerobatics_1	۲۱/۴	۲۶	۲۴	۲۸
aerobatics_2	۱۰/۹	۷/۷	۸/۹	۸/۲
big_1	۱۷/۸	۳۷/۴	۲۷/۱	۲۶/۱
big_2	۱۵/۸	۳۹	۳۴	۳۶
multiple_1	۹/۳	۱۲/۵	۱۳	۱۴/۳
multiple_2	۸/۹	۱۰/۹	۱۱/۷	۱۰
multiple_smoke_1	۱۲/۳	۱۴/۹	۱۰/۲	۱۱/۵
multiple_smoke_2	۷/۲	۱۶/۱	۱۴/۳	۱۵/۹
occlusion_1	۱۱/۱	۱۴	۱۷/۴	۲۳

جدول (۴). میزان هم‌پوشانی ناحیه هدف با خروجی حقیقی

Video sequence	ASIFT Tracker	[۱۷]	[۱]	[۱۸]
aerobatics_1	۰/۸۸	۰/۸۲	۰/۸۴	۰/۸۰
aerobatics_2	۰/۹۳	۰/۹۶	۰/۹۴	۰/۹۲
big_1	۰/۸۹	۰/۸۱	۰/۸۰	۰/۸۲
big_2	۰/۹۰	۰/۸۰	۰/۸۳	۰/۸۱
multiple_1	۰/۹۴	۰/۹۵	۰/۹۳	۰/۸۹
multiple_2	۰/۹۴	۰/۹۴	۰/۹۱	۰/۹۲
multiple_smoke_1	۰/۹۱	۰/۸۸	۰/۹۰	۰/۸۹
multiple_smoke_2	۰/۹۷	۰/۸۸	۰/۹۱	۰/۸۹
occlusion_1	۰/۹۵	۰/۹۲	۰/۸۷	۰/۸۳

۸- نتیجه‌گیری

در این مقاله روشی جدید برای ردیابی اشیاء صلب در دنباله‌های ویدئویی ارائه شد. در روش ارائه شده از الگوریتم ASIFT برای استخراج و تطبیق نقاط کلیدی، الگوریتم RANSAC برای حذف تناظرهای اشتباه، الگوریتم آشکارسازی اشیاء برجسته به منظور کنترل پوشیدگی و تاریخچه مقیاس سوژه برای مقابله با تغییر مقیاس هدف استفاده شد. الگوریتم ASIFT نسبت به شش پارامتر تبدیلات افاین مقاوم است؛ لذا در صورت تغییر زاویه دید، انتقال، دوران و تغییر مقیاس هدف نقاط تطبیقی با استفاده از این الگوریتم با دقت خوبی به دست می‌آید. در گام بعدی پس از استخراج نقاط کلیدی و یافتن نقاط تطبیقی توسط الگوریتم ASIFT، از الگوریتم RANSAC برای حذف تناظرهای اشتباه استفاده شد. هم‌چنین به منظور کنترل پوشیدگی نسبی یا پنهان

- Multiple Images of a Scene And some of Their Applications,” MIT press, 2001.
- [15] X. Hou and L. Zhang, “Saliency Detection: A spectral residual approach,” IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1-8.
- [16] H. Barlow, “Possible Principles Underlying the Transformation of Sensory Messages,” Sensory Communication, pp. 217-234, 1961.
- [17] M. Anvaripour, R. S. Kandovan, and S. Soltanpour, “Low Distance Airplanes Detection and Tracking Visually using Spectral Residual and KLT Composition,” Information Systems & Telecommunication, p. 159, 2014.
- [18] J. Sun, “A Fast MEANSHIFT Algorithm-Based Target Tracking System,” Sensors, vol. 12, pp. 8218-8235, 2012.
- [19] D. Wang, H. Lu, and M.-H. Yang, “Online object tracking with sparse prototypes,” IEEE Transactions on Image Processing, 2013, vol. 22, pp. 314-325.

Archive of SID

A New Framework for Tracking Aerial Moving Targets using AFFINE – SIFT Algorithm

V. Azadzadeh, A. M. latif*

*Yazd University

(Received: 07/06/2016, Accepted: 31/10/2016)

ABSTRACT

Abstract: Moving target tracking is a process in which an object is tracked and its location is determined in each frame. The goal of this process is facilitating the subsequent process in order to analyze the behavior or detect moving objects. In this paper, a new approach has been proposed for aerial moving targets detection and tracking based on feature matching algorithms. By this way, we propose spectral density for target detection and ASIFT feature matching algorithm for tracking. The challenge is selecting features that are robust against the changes of brightness, noise, rotation, scaling and viewing angle. To solve this problem, key points and their correspondence on the patterns extracted from consecutive frames, are calculated by the ASIFT algorithm. Also, to reduce false matches in consecutive frames, the RANSAC algorithm is used. In addition to strengthening the proposed algorithm against the scale change of target, the object history scale in the 10 previous frames is used. The proposed algorithm was performed on an AIRCRAFT TRACKING standard database. Experimental results demonstrate the robustness and accuracy of our proposed technique.

Keywords: Target tracking, ASIFT algorithm, RANSAC algorithm, Object history scale, Spectral density

* Corresponding Author Email: alatif@yazd.ac.ir