

# بهبود دهمی سرعت و مقاومت ردیابی تصویری تنگ مینا در مواجهه با انسداد جزئی به کمک مدل سازی

## زبر و نرم منظر هدف

محمد ترقی خواه<sup>۱</sup>، رسول امیرفتاحی<sup>۲</sup>، علیرضا معمارمقدم<sup>۳</sup>

۱ کارشناسی ارشد برق مخابرات، آزمایشگاه تحقیقاتی پردازش سیگنال، دانشکده مهندسی برق و کامپیوتر دانشگاه صنعتی اصفهان، m.taraghikah@ec.iut.ac.ir

۲ دانشیار دانشکده مهندسی برق و کامپیوتر، آزمایشگاه تحقیقاتی پردازش سیگنال، دانشگاه صنعتی اصفهان

۳ دانشجوی دکتری مهندسی برق (پردازش تصویر)، دانشکده فنی و مهندسی دانشگاه اصفهان

تاریخ دریافت: ۹۳/۸/۳۰ تاریخ پذیرش: ۹۵/۵/۲۷

## چکیده

در این مقاله الگوریتمی طراحی شده است که علاوه بر مقاوم بودن در برابر انسداد جزئی، حتی الامکان در برابر سایر چالش‌های ردیابی نیز عملکرد خوبی داشته باشد و همچنین از نظر زمانی، به حالت بلادرنگ نزدیک شود. برای این منظور ردیاب مبتنی بر نمایش تنک دو مرحله‌ای زبر و نرم را ارائه می‌دهیم که در آن ابتدا منظر شیء هدف با واژه‌نامه‌ای متشکل از بردارهای پایه PCA و قالب‌های جزئی مدل می‌شود و برای این مدل‌سازی از روش حل APG\_L1 بهره می‌بریم. برای افزایش سرعت محاسبات، در مرحله اول به کمک بلوک بندی مربعی و درشت روی قالب‌های جزئی، تعداد قالب‌های جزئی واژه‌نامه را کاهش می‌دهیم و آنگاه در مرحله بعد، برای پیدا کردن انسداد در سطح پیکسل، کاندیداهای برتر بدست آمده از مرحله قبل را با روش PCA\_L1 مورد ارزیابی قرار داده تا کاندیدای نهایی بدست آید. نتایج کمی مقایسه با پنج الگوریتم ردیابی غیر تنک و هفت الگوریتم تنک مینا بر روی دوازده دنباله تصویر نشان می‌دهد که ردیاب پیشنهادی با سرعت اجرای ۱۴٫۲ قاب بر ثانیه، منجر به میانگین خطای مرکزی ۱۰٫۲ و میانگین درصد هم پوشانی ۷۴٫۷۵ می‌شود؛ یعنی در مقایسه با شیوه‌های ردیابی متداول هم در زمینه دقت و هم در زمینه سرعت عملکرد خوبی داشته ایم و توانسته‌ایم بین عملکرد و بار محاسباتی ردیابی تعادل خوبی ایجاد کنیم.

## کلیدواژه

ردیابی تصویری، بیان تنک، تحلیل مؤلفه اصلی (PCA)، انسداد جزئی، مدل‌سازی منظر

## مقدمه

نیست؛ بنابراین منظر شیء هدف دچار تغییر شده که این تغییر اگر به درستی و به موقع شناسایی نشود، به دلیل رانش نتیجه، منجر به کاهش دقت ردیابی شده و در بدترین حالت موجب از دست رفتن هدف می‌شود. انسداد جزئی به سه صورت رخ می‌دهد: انسداد با شیء متحرک دیگر، انسداد با پس‌زمینه و انسداد شیء با خودش (در مورد اشیاء غیرصلب)؛ در این مقاله بیشتر متمرکز بر دو نوع اول هستیم.

معمولاً یک الگوریتم ردیابی شامل دو قسمت مهم می‌باشد [۱]: الف) مدل حرکت؛ در هر قاب برای شیء هدف تعدادی کاندیدا معرفی می‌کند مثل فیلتر کالمن یا فیلتر ذره‌ای<sup>۲</sup>. ب) مدل منظر؛ احتمال درست‌نمایی یک کاندیدا را ارزیابی می‌کند. از مشخصه‌های مهم یک الگوریتم کارآمد، داشتن مدل منظر قوی با قابلیت به‌روز رسانی می‌باشد. بیان تنک<sup>۳</sup> یکی از روش‌هایی است که دارای قابلیت به‌روز رسانی می‌باشد و برای مدیریت انسداد می‌توان از آن استفاده کرد؛ نخستین بار بیان تنک به‌وسیله می و

ردیابی تصویری، یکی از مهمترین مسائل در بینایی ماشین می‌باشد که هم به دلیل گستره وسیع کاربردهای آن نظیر تحلیل حرکت، تشخیص فعالیت، نظارت و مراقبت، تعامل انسان با رایانه و هم به دلیل چالش‌های فراوان ناشی از فاکتورهای ذاتی (همچون تغییر موقعیت و شکل شیء) و یا بیرونی (از قبیل تغییر روشنایی و نورپردازی محیطی، حرکت دوربین و مات‌شدگی صحنه، وجود هدف‌نما، اختلالات نویزی و انسداد)، همچنان توجه محققان را به خود جلب نموده است.

علی‌رغم پیشرفت‌های اساسی، انسداد جزئی همچنان مشکل اصلی در ردیابی تصویری می‌باشد. اگرچه برخی از ردیاب‌های ارائه شده در این زمینه، در مدیریت انسداد جزئی خوب عمل می‌کنند ولی اغلب آن‌ها در دنباله تصاویر چالش‌دار با تغییر منظر زیاد، دچار شکست می‌شوند. همچنین، الگوریتمی ارائه نشده است که بطور توأمان دارای سرعت و دقت کافی و مناسب باشد. موقعی که انسداد جزئی رخ می‌دهد، بخشی از شیء هدف قابل مشاهده

2 Particle Filter  
3 Sparse Representation

1 Occlusion

جزئی می‌باشد. سپس درست‌نمایی بر اساس ضرایب تجزیه محاسبه می‌شود که اندازه ضرایب، میزان شباهت بین قالب مرجع و کاندیداهای هدف را بیان می‌کنند. وانگ و همکاران [۹] در سال ۲۰۱۳ برای بیان بهتر شیء هدف، به جای سوابق زمانی استفاده از بردارهای پایه PCA را پیشنهاد داده‌اند. ایشان برای در نظر گرفتن اثر پیکسل‌های نویزی در هنگام ارزیابی یک کاندیدا و همچنین موقع به‌روز کردن پایه‌های PCA، از یک نقشه انسداد استفاده می‌کنند. جیا و همکاران [۱۰] یک مدل منظر تنک محلی ساختاری را ارائه کرده‌اند؛ در این روش ابتدا برای هر وصله محلی کاندیدای هدف یک بیان تنک بدست آورده و در ادامه به کمک یک روش ادغام همراستا<sup>۴</sup>، ضرایب تنک را با هم ترکیب کرده و ویژگی‌هایی را بدست می‌آورد که از آن‌ها برای تعیین کاندیدای برتر استفاده می‌کند. ژانگ و همکاران [۱۱] دو مدل مستقل مبتنی بر بیان تنک توسعه داده‌اند و برای ارزیابی کاندیداها از ترکیب اطلاعات هر دو استفاده می‌کنند. آقای ژوانگ و همکاران [۱] برای اینکه بتواند قالب‌های مثبت و منفی را از طریق یک مسئله بهینه‌سازی بر اساس کاندیداها بیان کند، مدل کمیته‌سازی ردیاب L1 را اصلاح کرده و سپس برای مدل اصلاح شده، یک روش حل جدید ارائه داده است. در ادامه و پس از تغییر ضرایب بدست آمده به کمک وزن‌های وفقی، از یک طرح ادغام جمعی استفاده کرده و امتیاز نهایی را برای هر کاندیدا بدست می‌آورد.

علاوه بر ردیاب‌های تنک، در ردیاب‌های غیر تنک نیز هنوز مشکل داشتن سرعت و دقت کافی در مواجهه با انسداد و سایر چالش‌ها وجود دارد؛ مثلاً در [۱۲] از یک زیرفضای PCA برای پرداختن به تغییرات منظر ناشی از تغییر روشنایی، چرخش و مقیاس استفاده شده ولی برای مدیریت انسداد چاره‌ای تدبیر نشده است. در [۱۳] برای مواجهه با انسداد از یک بیان قطعه-مینا مبتنی بر هیستوگرام رنگ استفاده می‌شود ولی به دلیل به‌روز نشدن قالب هدف، این روش به تغییرات منظر حساس است. برای جلوگیری از ایجاد رانش ناخواسته که به دلیل استفاده مستقیم از نتیجه ردیابی در به‌روز کردن ایجاد می‌شود، مرجع [۱۴] یک الگوریتم ردیابی برخط تقویتی پیشنهاد داده است که از یک ساختار آموزش چند نمونه‌ای کمک می‌گیرد. همچنین در [۱۵] برای آموزش کلاس‌بند براساس داده‌های برچسب‌دار و بدون برچسب از یک ایده نیمه نظارتی<sup>۵</sup> استفاده شده است. دو روش آخر، اگرچه در کاهش رانش دارای موفقیت خوبی هستند ولی در آن‌ها برای پرداختن مستقیم به انسداد تلاشی صورت نگرفته است.

هدف این مقاله، رسیدن به یک الگوریتم ردیابی می‌باشد که علاوه بر داشتن مقاومت خوب در برابر انسداد جزئی و چالش‌هایی از قبیل تغییر روشنایی و تغییر مقیاس، از نظر زمانی نیز معقول و به

همکاران [۲] در سال ۲۰۰۹ تحت عنوان "ردیاب L1" وارد ادبیات ردیابی شده و از آن زمان به بعد، بیان تنک در ساختار ردیاب‌های بسیاری با نام ردیاب‌های تنک بکار گرفته شده است. ایشان ردیابی را به صورت مسئله تقریب تنک در یک زیرفضای قالب (شامل سوابق زمانی شیء هدف و قالب‌های جزئی) با چارچوب کلی فیلتر ذره‌ای در نظر گرفته‌اند و برای مدیریت انسداد جزئی، استفاده از قالب‌های جزئی را پیشنهاد داده‌اند؛ حل تقریب تنک با یک مسئله کمیته‌سازی L1 انجام می‌گیرد. در ادامه درست‌نمایی هر کاندیدا با استفاده از خطای بدست آمده در بازسازی با سوابق زمانی شیء هدف محاسبه می‌شود. چون در این ردیاب، تقریب در هر قاب و روی همه کاندیداها باید اعمال شود، این ردیاب نیازمند حل یک مجموعه از مسائل تقریب تنک می‌باشد و بنابراین حجم محاسبات این ردیاب بسیار بالا است. برای پیدا کردن جواب‌های کارآمدتر در همین چارچوب، در سال ۲۰۱۱ نسخه کامل‌تری از این روش در [۳] مطرح شده که در آن برای کاهش تعداد ذراتی که نیاز است به صورت تنک تجزیه شوند یک جواب تقریبی توسعه داده شده است. برای سرعت بخشیدن به حل مسئله کمیته‌سازی L1، در [۴] از یک ایده کاهش گرادیان<sup>۵</sup> استفاده شده است که در آن برای بهینه کردن استفاده از قالب‌های جزئی، مدل کمیته‌سازی ردیاب L1 را اصلاح کرده و در ادامه برای حل مدل اصلاح شده، یک روش عددی سریع پیشنهاد شده است.

پس از بیان طرح ارزیابی کاندیداها توسط [۲] فرمول‌بندی‌های جدیدی در برخی ردیاب‌های تنک به‌وجود آمده‌اند. مثلاً ژانگ و همکاران [۵] در سال ۲۰۱۲ با استفاده از بیان تنک، ردیابی را به صورت یادگیری چندوظیفه‌ای<sup>۶</sup> فرمول‌بندی کرده‌اند. در این فرمول‌بندی، یادگیری بیان هر کاندیدای هدف به عنوان یک وظیفه در نظر گرفته می‌شود. در این روش با بیان توأم کاندیداها در یک یادگیری چند وظیفه‌ای، هم عملکرد و هم سرعت ردیاب بهبود می‌یابد. الگوریتم دیگری که روابط بین کاندیداها را در نظر می‌گیرد توسط ایشان در [۶] مطرح شده که در آن مسئله ردیابی به صورت یک مسئله آموزش ماتریس پایین-رتبه<sup>۷</sup> در نظر گرفته شده است.

اگرچه فرمول‌بندی‌های جدید مثل [۷]-[۵] در مدل‌سازی شیء هدف مؤثر هستند ولی طرح ارزیابی کاندیداها مبتنی بر خطای بازسازی، همواره کارآمد و مقاوم نیست. بنابراین چندین ردیاب مدل‌های تنک جدیدشان را به همراه اصلاح طرح ارزیابی کاندیدا بیان کرده‌اند؛ یکی از این روش‌ها توسط لیو و سان [۸] ارائه شده که در آن قالب مرجع را بر اساس کاندیداها هدف و قالب‌های جزئی بیان می‌کند یعنی واژه‌نامه متشکل از همه کاندیداها و قالب‌های

4 Trivial Templates  
5 Gradient Descent  
6 Multi Task  
7 Low-Rank

8 Alignment Pooling  
9 Semi-Supervised

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = N(\mathbf{x}_t; \mathbf{x}_{t-1}, \boldsymbol{\varphi}) \quad (2)$$

$\boldsymbol{\varphi}$  ماتریس کواریانس قطری است (چون پارامترهای مستوی مستقل فرض شده‌اند). بنابراین برای تخمین حالت در قاب آتی، از یک توزیع نرمال با میانگین حالت در قاب جاری و با واریانس پارامترهای مستوی استفاده می‌شود.

### مدل منظر شیء مبتنی بر واژه‌نامه تصاویر ویژه

در این مدل از بردارهای پایه PCA به جای سوابق زمانی برای پرداختن به تغییرات منظر هدف استفاده می‌کنیم؛ سوابق زمانی بکار رفته در مدل منظر ردیاب L1، فقط نمونه‌هایی تصادفی از زیرفضای شیء هدف هستند که بسیار هم با یکدیگر دارای همبستگی می‌باشند و بنابراین دارای اطلاعات تکراری بسیاری هستند که برای بیان همه زیرفضای شیء هدف مناسب نیستند؛ ترکیب آن‌ها فقط اشیائی را تقریب می‌زند که بسیار شبیه خودشان هستند [16]. برای استفاده بهتر از بیان زیرفضا، منظر شیء هدف را می‌توان با  $k$  بردار پایه PCA مدل کرد؛ ترکیب خطی آن‌ها می‌تواند به خوبی هر کاندیدای خراب نشده از زیرفضای شیء هدف را بیان کند. پس از بدست آوردن قالب هر کاندیدای هدف ( $\mathbf{y}$ ) با توجه مدل حرکت، ابتدا میانگین آن را به صورت زیر صفر می‌کنیم:

$$\bar{\mathbf{y}} = \mathbf{y} - \boldsymbol{\mu} \quad (3)$$

که  $\boldsymbol{\mu}$  مرکز زیر فضای شیء هدف است. سپس کاندیدای با میانگین صفر را مورد مدل‌سازی قرار می‌دهیم؛ توجه کنید در هر دو مرحله پیشنهادی منظور از بردار مشاهده، کاندیدای هدف با میانگین صفر یعنی  $\bar{\mathbf{y}}$  است. در ادامه دو مرحله مدل‌سازی را بیان خواهیم کرد.

### مرحله اول: مدل‌سازی زیر منظر

یکی از روش‌هایی که برای افزایش سرعت حل پیشنهاد شده است، یک حل‌کننده عددی سریع بر اساس ایده APG می‌باشد؛ این روش برای حل مسئله تنک از طریق کمینه‌سازی معادله (4) ارائه شده است [4]:

$$\min_{\mathbf{c}} \frac{1}{2} \|\mathbf{y} - \mathbf{Dc}\|_2^2 + \lambda \|\mathbf{c}\|_1 + \frac{\mu_t}{2} \|\mathbf{e}\|_2^2 \quad \text{s.t. } \mathbf{z} \geq 0 \quad (4)$$

که در این رابطه  $\mathbf{y} \in \mathbb{R}^{d \times 1}$  بردار مشاهده،  $\mathbf{D} \in \mathbb{R}^{d \times (k+n)}$  ماتریس واژه‌نامه ( $k$  تعداد بردارهای پایه PCA و  $n$  تعداد قالب‌های جزئی است)،  $\mathbf{c} = [\mathbf{z}; \mathbf{e}]$  بردار ضرایب که به ترتیب شامل ضرایب قالب‌های هدف و قالب‌های جزئی است،  $\lambda$  پارامتر تنظیم و  $\mu_t$  پارامتری جهت کنترل انرژی ضرایب جزئی است که اگر آن را برابر صفر قرار دهیم به همان معادله کمینه‌سازی L1 خواهیم رسید [4]. در مرحله اول، ماتریس واژه‌نامه را به صورت زیر تشکیل می‌دهیم:

حالت بلادرنگ نزدیک شود؛ برای این منظور، الگوریتمی دو مرحله‌ای و مبتنی بر بیان تنک با حل سریع مسئله ارائه می‌کنیم. در ادامه مقاله، ابتدا الگوریتم پیشنهادی را مطرح نموده و سپس عملکرد آن را به لحاظ معیارهای شناخته شده کمی و کیفی مورد ارزیابی قرار می‌دهیم و در انتها جمع‌بندی خواهیم کرد.

### روش پیشنهادی

روش پیشنهادی، نسخه تصحیح یافته‌ای از روش OSPT [9] است؛ تفاوت دو روش در نحوه مدل‌سازی منظر شیء هدف می‌باشد. به منظور کاهش حجم محاسباتی ردیاب OSPT در کنار حفظ و یا حتی بهبود دقت، ایده زیر و نرم  $l^1$  را در مدل‌سازی منظر مطرح کرده‌ایم. بر این اساس، ردیابی در دو مرحله محقق می‌شود؛ در مرحله اول، برای افزایش سرعت محاسبات به ازای هر بلوک  $a \times a$  پیکسل از پنجره پیکسلی هدف، یک قالب جزئی در نظر می‌گیریم که با این ایده و همچنین بکارگیری روش عددی سریع [4] APG\_L1 برای مدل‌سازی، با سرعت خوبی به یک پاسخ اولیه می‌رسیم. نقص مرحله اول این است که در ازای افزایش سرعت محاسبات، دقت نتیجه را کاهش می‌دهد چون دیگر خطا را در سطح پیکسل مورد ارزیابی قرار نمی‌دهد. سپس در مرحله دوم برای افزایش دقت در تعیین هدف، کاندیداهای محتمل‌تر هدف که در مرحله اول شناسایی و گزینش شده‌اند را مورد ارزیابی قرار داده و به خطا در سطح پیکسل رسیدگی می‌کنیم؛ در این مرحله واژه‌نامه را با قالب‌های جزئی و بردارهای پایه PCA پُر کرده و برای مدل‌سازی از روش حل [9] PCA\_L1 استفاده می‌کنیم. مدل منظر در طول ردیابی با چارچوب فیلتر ذره‌ای، به صورت برخط به روز رسانی می‌شود. همانند روش OSPT به منظور کاهش انحراف از هدف، شیوه‌ای جهت احتساب انسداد بکارگرفته می‌شود. در ادامه مهم‌ترین مراحل روش پیشنهادی را بیان خواهیم کرد.

### مدل‌سازی حرکت در چارچوب فیلتر ذره‌ای

برای مدل کردن حرکت بین دو قاب متوالی از فرض تبدیل مستوی (آفاین<sup>11</sup>) استفاده می‌کنیم. شش پارامتر این تبدیل برای مدل کردن انتقال حالت شیء بکار گرفته می‌شوند؛ چنانچه حالت را به صورت زیر در نظر بگیریم:

$$\mathbf{x}_t = \{x_t, y_t, \theta_t, s_t, \alpha_t, \phi_t\} \quad (1)$$

که عناصر آن به ترتیب معرف جابجایی در جهات  $x$  و  $y$ ، زاویه چرخش، مقیاس، تناسب و اعوجاج قالب منظر شیء هستند؛ آنگاه انتقال حالت با قدم‌زدن تصادفی و توزیع نرمال فرمول‌بندی می‌شود یعنی:

10 Coarse and Fine  
11 Affine

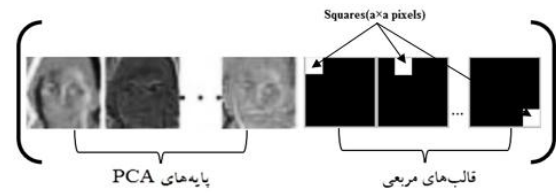
چنانچه خطای  $e_j^i$  پیکسل  $j$  ام قالب هدف برابر صفر باشد آنگاه  $w_j^i = 1$  و در غیر اینصورت  $w_j^i = 0$  می باشد.  $\square$  نشان دهنده ضرب هادامارد (عنصر در عنصر متناظر) و  $\beta$  ضریب جریمه است (در روش پیشنهادی  $\beta$  را برابر  $\lambda$  قرار می دهیم). قسمت اول عبارت نمایی به خطای بازسازی بخش های غیرانسدادی می پردازد و هدف قسمت دوم، جریمه کردن برچسب گذاری هر پیکسل به عنوان انسداد است.

### بهروز کردن مدل منظر و بردار میانگین

بهروز کردن مدل برای پرداختن به تغییرات منظر شیء هدف، امری ضروری است. استفاده از همه ی کاندیداهای برتر هدف، بدون تأثیر دادن پیکسل های انسدادی، موجب ایجاد و تشدید انحراف در ردیابی می شود. همانند روش OSPT از یک نقشه انسداد برای اثر دادن پیکسل های نویزی در فرآیند بهروز رسانی استفاده می کنیم؛ المان های غیر صفر این نقشه، بیان کننده ی پیکسل هایی است که دچار انسداد شده اند. برای این که بفهمیم از کاندیدای برتر قاب جاری برای بهروز رسانی استفاده کنیم و یا نه، نرخ تصرف انسداد ( $\eta$ ) را به شکل نسبت تعداد پیکسل های غیر صفر به تعداد کل پیکسل های نقشه انسداد محاسبه می کنیم. سپس دو آستانه  $tr_1$  و  $tr_2$  را برای توصیف درجه انسداد بکار می بریم. در آخر بر اساس مقدار  $\eta$  یکی از سه مورد بهروز رسانی کامل، بهروز رسانی جزئی و یا عدم بهروز رسانی را انجام می دهیم؛ اگر  $\eta < tr_1$  باشد، ما مدل را به صورت مستقیم با این نمونه بهروز می کنیم. اگر  $tr_1 \leq \eta \leq tr_2$  باشد، می فهمیم که هدف دچار انسداد جزئی شده و پیکسل های انسدادی را با پیکسل های مشابه از بردار میانگین جایگزین کرده و سپس از این نمونه اصلاح شده برای بهروز رسانی استفاده می کنیم. اگر  $\eta \geq tr_2$  باشد، این معنی را می دهد که بخش قابل توجهی از هدف دچار انسداد شده و بنابراین از این نمونه برای بهروز رسانی استفاده نمی کنیم (در روش پیشنهادی  $tr_1=0.1$  و  $tr_2=0.6$  در نظر گرفته شده است). پس از اینکه به تعداد کافی نمونه برای بهروز رسانی جمع کردیم، از روش [12] IPCA برای بهروز کردن بردارهای پایه PCA و بردار میانگین استفاده می کنیم.

### نتایج شبیه سازی

تمامی آزمایش های روش پیشنهادی توسط یک رایانه با پردازشگر پنج هسته ای ۲/۵ GHZ با حافظه ۴ GB و تحت نرم افزار MATLAB 2013a انجام گرفته است. دنباله تصاویر استاندارد مربوط به مراجع [۱۷] و [۱۸] می باشند. برای هر دنباله



شکل ۱. واژه نامه مرحله اول مدل سازی منظر

چون بُعد واژه نامه را کاهش داده ایم، حل ایده APG\_L1 با سرعت خوبی انجام می گیرد. عیب مرحله ی اول پیشنهادی این است که این روش انسداد جزئی و سایر عوامل چالش دار را در سطح پیکسل مورد بررسی قرار نمی دهد و در نتیجه مقاومت ردیاب را در برابر انسداد جزئی و سایر عوامل چالش دار تا حدودی کاهش می دهد. برای مدیریت انسداد در سطح پیکسل و همچنین بهبود دهی مقاومت ردیاب در برابر سایر عوامل چالش زا، استفاده از مرحله دوم را پیشنهاد می دهیم.

### مرحله دوم: مدل سازی نرم منظر

در این مرحله فقط کاندیداهای برتر (مثلاً ده درصد برتر) به دست آمده از مرحله اول مورد ارزیابی قرار می گیرند. این بار، مسئله کمینه سازی را به صورت زیر در نظر می گیریم:

$$\min_{z, e} \frac{1}{2} \|y - Uz - e\|_2^2 + \lambda \|e\|_1 \quad (5)$$

در رابطه بالا،  $U \in \mathbb{R}^{d \times k}$  ماتریسی از بردارهای پایه PCA است. در این مرحله از نمونه های مرجع تنک برای مدل سازی منظر هدف استفاده می کنیم؛ منظور از نمونه های مرجع تنک، تعداد کمی از بردارهای پایه PCA و مجموعه ای از قالب های جزئی است:



شکل ۲. نمونه های مرجع تنک با قالب های جزئی کامل

همچنانکه اشاره شد در این مرحله برای حل مسئله کمینه سازی و به دست آوردن ضرایب از روش PCA\_L1 استفاده می کنیم.

### شباهت سنجی کاندیداهای هدف با احتساب نقشه انسداد

پس از محاسبه ی ضرایب بیان کننده همه ی کاندیداهای میانگین صفر، حال می خواهیم معیار ارزیابی کاندیداها را بیان کنیم؛ در این معیار، اثر پیکسل های انسدادی وارد شده است:

$$p(\mathcal{J}^i | \mathcal{X}^i) = \exp[-(\|\mathcal{W}^i \square (\mathcal{J}^i - U\mathcal{Z}^i)\|_2^2 + \beta \sum (1 - \mathcal{W}^i))] \quad (6)$$

که  $z_i$  و  $e_i$  به ترتیب معرف ضرایب بردارهای پایه PCA و قالب های جزئی کاندیدا  $i$  ام هستند؛

## نتیجہ گیری

هدف این مقاله، رسیدن به یک الگوریتم ردیابی است که علاوه بر داشتن مقاومت خوب در برابر انسداد جزئی و برخی چالش‌ها از قبیل تغییر روشنایی، تغییر مقیاس و چرخش، از نظر بار محاسباتی نیز به حالت بلادرنگ نزدیک شود. در این راستا الگوریتمی دو مرحله‌ای و مبتنی بر بیان تنک ارائه داده‌ایم که در آن برای مدل‌سازی هدف و نیز تشخیص نرخ انسداد جزئی، ایده زبر و نرم را در تشکیل واژه‌نامه بیان تنک بکار برده‌ایم. نتایج ارزیابی کمی و کیفی نشان می‌دهند که روش پیشنهادی در مقایسه با سایر الگوریتم‌های ردیابی متداول، در مجموع دارای خطای کمتر و سرعت اجرای بیشتری می‌باشد.

## مرجع‌ها

- [1]Bohan Zhuang, Huchuan Lu, Ziyang Xiao and Dong Wang, "Visual tracking via discriminative sparse similarity map", TIP, vol. 23, no. 4, pp. 1872-1881, 2014.
- [2]X. Mei and H. Ling, "Robust Visual Tracking Using L1 Minimization", ICCV, pp. 1436-1443, 2009.
- [3]X. Mei, H. Ling, Y. Wu, E. Blasch and L. Bai, "Minimum Error Bounded Efficient L1 Tracker With Occlusion Detection", CVPR, pp. 1257-1264, 2011.
- [4]C. Bao, Y. Wu, H. Ling and H. Ji, "Real Time Robust L1 Tracker Using Accelerated Proximal Gradient Approach", CVPR, pp. 1830-1837, 2012.
- [5]T. Zhang, B. Ghanem, S. Liu and N. Ahuja, "Robust Visual Tracking Via Multi-Task Sparse Learning", CVPR, pp. 2042-2049, 2012.
- [6]T. Zhang, B. Ghanem, S. Liu and N. Ahuja, "Low-Rank Sparse Learning For Robust Visual Tracking", ECCV, pp. 470-484, 2012.
- [7]B. Liu, J. Huang, C. Kulikowski, L. Yang and C. Kulikowski, "Robust Tracking Using Local Sparse Appearance Model And K-Selection", CVPR, vol. 35, no. 12, pp. 1313-1320, 2011.
- [8]H. Liu and F. Sun, "Visual Tracking Using Sparsity Induced Similarity", ICPR, pp. 1702-1705, 2010.
- [9]D. Wang, H. Lu and M.-H. Yang, "Online Object Tracking With Sparse Prototypes", TIP, vol. 22, no. 1, pp. 314-325, 2013.
- [10]X. Jia, H. Lu and M. Yang, "Visual Tracking Via Adaptive Structural Local Sparse Appearance Model", CVPR, pp. 1822-1829, 2012.
- [11]W. Zhong, H. Lu and M. Yang, "Robust Object Tracking Via Sparsity-Based Collaborative Model", CVPR, pp. 1838-1845, 2012.
- [12]D. A. Ross, J. Lim, R.-S. Lin and M.-H. Yang, "Incremental Learning For Robust Visual Tracking", IJCV, vol. 77, no. 1-3, pp. 125-141, 2008.
- [13]A. Adam, E. Rivlin and I. Shimshoni, "Robust Fragments-based Tracking using the Integral Histogram", CVPR, vol. 1, pp. 798-805, 2006.
- [14]B. Babenko, M.-H. Yang and S. Belongie, "Visual Tracking with Online Multiple Instance Learning", TPAMI, pp. 983-990, 2010.
- [15]Z. Kalal, J. Matas and K. Mikolajczyk, "P-N Learning: Bootstrapping Binary Classifiers By Structural Constraints", CVPR, pp. 49-56, 2010.

تصاویر، حالت شیء هدف در قاب اول به صورت دستی مشخص شده است. در بیان PCA، هر مشاهده تصویر به یک کاندیدای هدف با ابعاد  $32 \times 32$  پیکسل نرمالیزه شده و از ۱۶ بردار پایه استفاده کرده‌ایم که به صورت افزایشی و هر ۵ قاب یکبار به روز می‌شوند. برای ایجاد بده-بستانی بین عملکرد و بار محاسباتی، برای قاب آتی ۲۰۰ کاندیدای هدف تخمین زده می‌شود. در مرحله اول  $a = 8$ ،  $\lambda = 0.01$  و  $\mu_t = 0$  در نظر گرفته شده و در مرحله دوم  $\lambda = 0.05$  قرار داده می‌شود. روش پیشنهادی با ردیاب‌های مرسوم غیرتنک [12] IVT، [15] PN، [19] VTD، [14] MILTrack، [13] FragTrack و همچنین ردیاب‌های تنک [4] APGL، [5] MTT، [10] ASLAS، [7] LSAT، [11] SCM، [9] OSPT و [1] DSS مورد مقایسه قرار گرفته است. عملکرد روش پیشنهادی در مواجهه با چالش‌های مختلف، روی مجموعه‌ای شامل دوازده دنباله تصاویر مرسوم دارای تغییرات منظری از قبیل انسداد، تغییر روشنایی، تغییر مقیاس، چرخش و سایر چالش‌ها از طریق ارزیابی کیفی و کمی با سایر روش‌ها مورد مقایسه قرار گرفته است. با توجه به اینکه دسترسی به مجموعه داده‌های اندازه‌گیری شده ردیاب‌های تنک امکان‌پذیر نبود، ارزیابی کیفی تنها بر روی ردیاب‌های غیرتنک انجام شده ولی در ارزیابی کمی، مقایسه بین همه ردیاب‌ها صورت می‌گیرد.

## ارزیابی کیفی

نمونه قاب‌هایی از نتیجه‌ی اجرای ردیابی بر روی دنباله تصاویر مختلف در شکل (۳) نمایش داده شده‌اند. همانطور که از روی شکل نیز مشخص است، ردیاب پیشنهادی در مواجهه با چالش‌های مختلف، در مجموع دارای عملکرد بصری بهتری است؛ یعنی مستطیل محصور کننده هدف بدست آمده توسط ردیاب پیشنهادی، دارای همپوشانی بسیار خوبی با هدف در نمونه قاب‌های نمایش داده شده است.

## ارزیابی کمی

در جدول (۱) روش‌های مختلف با معیارهای ارزیابی خطای مرکزیابی و درصد هم‌پوشانی مورد مقایسه قرار گرفته‌اند. در هر سلول این جدول، عدد سطر اول مربوط به خطای مرکزیابی و عدد سطر دوم مربوط به درصد هم‌پوشانی می‌باشد. در هر ردیف، سه نتیجه برتر یعنی سه ردیاب دارای خطای کمتر در مورد هر دنباله تصاویر به ترتیب با رنگ‌های قرمز، آبی و سبز نشان داده شده‌اند. دو ردیف آخر این جدول نرخ قاب بر ثانیه (fps) را نشان می‌دهند که یکی با احتساب زمان نمایش نتیجه بر روی قاب و دیگری بدون محاسبه این زمان بدست آمده است. با توجه به این جدول، ردیاب پیشنهادی هم به لحاظ دقت و هم سرعت دارای بهترین عملکرد است؛ بنابراین توانسته‌ایم یک تعادل خوب بین عملکرد و بار محاسباتی ایجاد کنیم.



- [16]Z. Xiao, H. Lu and D. Wang, "Object Tracking with L2-RLS", *ICPR*, pp. 1351–1354, 2012.  
 [17]<http://faculty.ucmerced.edu/mhyang/pubs.html>.  
 [18]<http://ice.dlut.edu.cn/lu/publications.html>.  
 [19]J. Kwon and K. M. Lee, "Visual Tracking Decomposition", *CVPR*, pp. 1269–1276, 2010.



(الف)



(ب)



(ج)

— IVT — PN — VTD — MIL — Frag — OSPT — Our Tracker

شکل ۳. ارزیابی کیفی: (الف) انسداد و تغییر مقیاس (ب) تغییر روشنایی و چرخش (ج) حرکت سریع و هدف‌نما

جدول ۱. ارزیابی کمی: در هر سلول، عدد سطر اول مربوط به خطای مرکزی و عدد سطر دوم مربوط به درصد هم پوشانی می باشد.

Methods Sequences	IVT	PN	VTD	MIL	Frag	APGL1	MTT	ASLAS	LSAT	SCM	OSPT	DSS	Our Tracker
Occlusion1	9.2 0.85	17.5 0.65	11.1 0.77	32.3 0.59	5.6 0.90	6.8 0.87	14.1 0.79	10.8 0.83	5.3 0.90	3.2 0.93	4.7 0.91	4.5 0.91	4.7 0.91
Occlusion2	10.2 0.59	18.6 0.49	10.4 0.59	14.1 0.61	15.5 0.60	-	-	-	-	-	4 0.84	6.72 0.76	5.57 0.78
Caviar2	8.6 0.45	8.5 0.66	4.7 0.67	70.3 0.25	5.6 0.56	63.1 0.32	65.4 0.33	62.3 0.35	45.6 0.28	2.5 0.81	2.2 0.71	3.6 0.77	2 0.73
Deer	127.6 0.22	25.7 0.41	11.9 0.58	66.5 0.21	92.1 0.08	38.4 0.45	9.2 0.60	8 0.62	69.8 0.35	36.8 0.46	8.5 0.61	8.8 0.63	13.73 0.55
Jumping	36.8 0.28	3.6 0.69	63 0.08	9.9 0.53	58.5 0.14	8.8 0.59	19.2 0.30	39.1 0.24	55.2 0.09	3.9 0.73	5 0.69	6 0.66	4.66 0.69
Singer1	8.5 0.66	32.7 0.41	4.1 0.79	15.2 0.34	22 0.34	3.1 0.83	41.2 0.32	5.3 0.78	14.5 0.52	3.7 0.85	4.7 0.82	3.9 0.85	4.2 0.83
David Indoor	3.6 0.71	9.7 0.60	13.6 0.52	16.2 0.45	76.7 0.19	14.3 0.57	124 0.28	6 0.63	4.9 0.62	5 0.69	3.7 0.80	3.3 0.80	3.48 0.79
David Outdoor	53 0.52	173 0.16	61.9 0.42	38.4 0.41	90.5 0.39	-	-	-	-	-	5.8 0.77	99 0.33	50.6 0.58
Car4	2.9 0.92	18.8 0.64	12.3 0.73	60.1 0.34	179.8 0.22	16.4 0.70	37.2 0.53	4.3 0.89	3.3 0.91	3.5 0.89	3 0.92	2.4 0.92	2.66 0.92
Car11	2.1 0.81	25.1 0.38	27.1 0.43	43.5 0.17	63.9 0.09	1.7 0.83	1.8 0.58	2 0.81	4.1 0.49	1.8 0.79	2.2 0.81	1.8 0.79	1.42 0.85
Face	69.7 0.44	22.3 0.62	141.4 0.24	134.7 0.15	48.8 0.39	57.8 0.35	127.2 0.26	95.1 0.21	16.5 0.69	125.1 0.36	14.95 0.70	8.5 0.77	14.49 0.71
Sylvester2008b	70.8 0.51	16.8 0.59	49.2 0.44	31.3 0.54	98.9 0.05	112.4 0.28	14.6 0.63	9.1 0.74	14.4 0.73	9 0.65	99.6 0.13	6.7 0.76	15.1 0.63
Average	33.6 0.58	31 0.53	34.2 0.52	44.4 0.38	63.1 0.33	32.3 0.58	45.4 0.46	24.2 0.61	23.4 0.56	19.5 0.72	13.2 0.73	12.9 0.7458	10.2 0.7475
fps(with showing results)	32	18	4	32	4	4	1	9	2	0.5	2	2	10.4
fps(without showing results)	-	-	-	-	-	-	-	-	-	-	2.7	2.7	14.2

