

بسم الله الرحمن الرحيم



نشریه سامانه‌های غیرخطی در مهندسی برق

دوره ۱۲، شماره ۱، بهار و تابستان ۱۴۰۴

JNSEE
سامانه‌های غیرخطی در مهندسی برق



ISSN: 2322-3146

صاحب امتیاز: دانشگاه صنعتی سهند - معاونت پژوهش و فناوری

مدیر مسئول: دکتر رضا فلاح

سرمدیر: دکتر علی بهرامی

مدیر اجرایی: خانم دکتر آرزو جهانی

ناشر: دانشگاه صنعتی سهند با همکاری مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری و پایگاه استنادی علوم جهان اسلام (ISC)

حروفچینی و صفحه‌آرایی: دکتر آرزو جهانی

هیأت تحریریه:

دکتر بهنام محمدی ایواتلو - استاد گروه قدرت دانشگاه LUT فنلاند

دکتر فرزاد هاشم زاده - استاد گروه کنترل دانشگاه ITU ترکیه

دکتر علی رستمی - استاد گروه الکترونیک دانشگاه تبریز

دکتر فرخ مروستی - استاد گروه مخابرات دانشگاه صنعتی شریف

دکتر آرش امینی - استاد گروه مخابرات دانشگاه صنعتی شریف

دکتر محمداقبر شمس اللهی - استاد گروه مهندسی پزشکی دانشگاه صنعتی شریف

دکتر افشین ابراهیمی - استاد گروه مخابرات دانشگاه صنعتی سهند تبریز

دکتر اسماعیل نجفی اقدم، استاد گروه الکترونیک دانشگاه صنعتی سهند تبریز

دکتر اکبر غفارپور رهبر - استاد گروه مهندسی کامپیوتر دانشگاه صنعتی سهند تبریز

دکتر علی بهرامی - استاد گروه الکترونیک دانشگاه صنعتی سهند تبریز

گروه ویراستاری: دکتر رسول کاظم زاده (قدرت)، دکتر رضا محبوبی (کنترل) و دکتر معصومه آذغانی (الکترونیک و مخابرات)

نشریه سامانه‌های غیرخطی در مهندسی برق با مجوز شماره ۱۵۹۸۱۱/۳/۱۸ مورخ ۱۳۹۱/۰۸/۰۲ و کسب اعتبار علمی-پژوهشی با مجوز شماره ۱۰۲۶۱/۱۸/۳ مورخ ۱۳۹۶/۰۵/۱۵ از کمیسیون نشریات علمی-پژوهشی وزارت علوم، تحقیقات و فناوری به زبان فارسی منتشر شده و در پایگاه استنادی جهان اسلام (ISC) نمایه می‌شود. این نشریه تخصصی در راستای تحقق اهداف و توسعه مرزهای دانش و به منظور گسترش و نشر پژوهش‌های بنیادی، کاربردی و توسعه‌ای، مقالات اصیل علمی در گرایش‌های مختلف مهندسی برق و کامپیوتر که قبلاً در هیچ نشریه دیگری انتشار نیافته باشند را با تأکید بر تحلیل و طراحی سامانه‌های غیرخطی به چاپ می‌رساند. مطالب مندرج در مقالات صرفاً بیانگر نقطه نظرات نویسندگان بوده و این آراء لزوماً نظر مسئولین نشریه نیستند.

نشانی پستی نشریه: تبریز، شهر جدید سهند، پردیس دانشگاه صنعتی سهند، معاونت پژوهش و فناوری، دبیرخانه نشریات تخصصی دانشگاه،

کد پستی: ۵۳۳۱۸/۱۷۶۳۴

تلفن: ۰۴۱-۳۳۴۵۹۳۴۰ و ۰۴۱-۳۳۴۵۹۳۲۱ دورنگار: ۰۴۱-۳۳۴۴۳۸۷۷

پست الکترونیک: jnsee@sut.ac.ir سایت اینترنتی: <http://jnsee.sut.ac.ir>

نشریه سامانه‌های غیرخطی در مهندسی برق

دوره ۱۲، شماره ۱، بهار و تابستان ۱۴۰۴

فهرست مقالات

شماره صفحه	عنوان
۵	یک روش کارآمد برای تناظریابی اشکال سه‌بعدی با استفاده از ویژگی‌های هندسی نویسندگان: امیررضا امیرفتحیان، حسین ابراهیم نژاد
۳۶	تشخیص رفتارهای غیر طبیعی رانندگان از طریق آموزش الگوهای باینری محلی و هیستوگرام گرادیان جهت‌دار نویسندگان: نیما کاوه مریان، حسین ابراهیم نژاد، احمدرضا یگانه مهر
۵۷	موقعیت‌یابی ربات متحرک در محیط‌های داخلی با استفاده از ترکیب فیلتر کالمن توسعه یافته و الگوریتم مسیریابی درخت تصادفی نویسندگان: محمدرسول حاج علی، رمضان هاونگی
۸۱	کنترل مقید عمق یک رونده زیرسطحی مبتنی بر کنترل‌کننده پیش‌بین تعمیم‌یافته: پیاده‌سازی پردازنده در حلقه نویسندگان: سیدحسین میرباقری، سعید نصرالهی
۱۰۰	تشخیص وسیله نقلیه از تصاویر با بهره‌گیری از الگوریتم تغییر یافته YOLOv4 با دقت بالا در یادگیری عمیق نویسندگان: جواد صیادی، مهدی نان‌گیر، بهزاد مظفری تازه‌کند، محمودمحصل فقهی، حمید صیادی
۱۱۸	کنترل پیش‌بین مدل تطبیقی برای مساله کنترل بار-فرکانس سیستم‌های قدرت بهم‌پیوسته چند ناحیه‌ای با پارامترهای متغیر با زمان نویسندگان: امین خادم، محسن احمدنیا و احمد حاجی‌پور

Journal of Nonlinear Systems in Electrical Engineering

Vol.12, No.1, Spring and Summer 2025

List of articles

Title	Page
An efficient method for matching 3D shapes using geometric features Authors: Amirreza Amirfathiyani, Hossein Ebrahimnezhad	5
Detection of Abnormal Driver Behaviors Using Local Binary Patterns and Histogram of Oriented Gradients Authors: Nima Kaveh Maryan, Hossein Ebrahimnezhad, Ahmadreza Yeganeh Meh	36
Mobile Robot Localization in Indoor Environments Using a Hybrid Extended Kalman Filter and Rapidly-Exploring Random Tree Algorithm Authors: Mohammadrasoul Haj Ali, Ramazan Havangi	57
Depth-Constrained Control of an Autonomous Underwater Vehicle Based on a Generalized Model Predictive Controller: Processor-in-the-Loop Implementation Authors: Seyed Hosein Mirbagheri Saeed Nasrollahi	81
Vehicle Detection from Images Using a Modified YOLOv4 Algorithm with High Accuracy in Deep Learning Authors: Javad Sayyadi, Mahdi Nangir, Behzad Mozaffari Tazekand, Mahmood Mohassel Fegghi, Hamid sayyadi	100
Adaptive Model Predictive Control for Load Frequency Control Problem in Multi-Area Interconnected Power Systems with Time-Varying Parameters Authors: Amin Khadem, Mohsen Ahmadnia, Ahmad Hajipour	118



Sahand University
of Technology

DOR:

[20.1001.1.23223146.1404.12.1.1.4](https://doi.org/10.1001.1.23223146.1404.12.1.1.4)

Journal of Nonlinear
Systems in Electrical
Engineering

Vol.12, No.1

Spring and Summer 2025

ISSN: 2322 – 3146

<http://jnsee.sut.ac.ir>

An efficient method for matching 3D shapes using geometric features

Amirreza Amirfathiyan, Hossein Ebrahimnezhad*

Computer Vision Res. Lab., Faculty of Electrical Engineering, Sahand University of Technology, Tabriz, Iran.

E-mails: a_amirfathiyan@sut.ac.ir; ebrahimnezhad@sut.ac.ir

Abstract

In this article, an efficient method for finding the correspondence between 3D models containing dynamic and non-rigid deformations is proposed, which uses the geometric features of the models. For this purpose, an intrinsic feature based on geodesic distances of points is used and the effectiveness of the proposed method for matching three-dimensional shapes is shown. First, a set of landmarks (salient points or special points) are identified and selected by sampling the farthest point. Then, the feature matrix of the models is created based on the geodesic distance of the points of the 3D shapes from the landmarks. In the following, a linear assignment problem is solved by applying the LAPJV algorithm, whose cost function contains the feature matrix of the models. Conducting extensive tests on different models from the standard data set shows the quantitative and qualitative superiority of the proposed method compared to other methods including approaches [48], [50] and [60]. The results show that the proposed method achieves better results in accuracy and efficiency compared to the methods [48], [50] and [60] on models with dynamic and non-rigid deformations and non-isometric cases.

Keywords

3D Shape Matching; Salient point Detector; Geodesic distance descriptor; Linear Assignment Problems using LAPJV Algorithm; Non-isometric deformations

1. Short Introduction

Finding correspondence between 3D shapes is a fundamental task in geometric processing, with applications in shape interpolation, reconstruction, and deformation transfer. Traditional methods, such as those relying on heat kernel signatures or wave kernel maps, excel in isometric deformations but struggle with non-rigid, motional, and non-isometric changes due to altered geodesic distances and computational complexity. This paper proposes an efficient approach using intrinsic geometric features based on geodesic distances from salient points detected via farthest point sampling. The feature matrices are matched through a linear assignment problem solved by the LAPJV algorithm, enhancing accuracy and efficiency. Motivated by the need to address real-world challenges like severe pose variations and non-isometric deformations in standard datasets, our method demonstrates superior quantitative and qualitative performance over existing techniques.

2. Proposed Work

The proposed method designs an efficient framework for 3D shape correspondence, comprising three key stages. First, salient points are detected using farthest point sampling based on geodesic distances, approximated via average geodesic distance (AGD) and Fast Marching algorithm, with 16 points selected to balance accuracy and computation. Second, feature matrices are extracted by computing normalized geodesic distances from all vertices to these salient points, forming a robust geometric descriptor. Third, correspondence is established by solving a linear assignment problem (LAP) using the LAPJV algorithm, where the cost matrix derives from Euclidean differences between feature matrices, ensuring unique point-to-point mappings with low complexity. Experimental evaluations were conducted on DT4D (non-rigid deformations) and SMAL (non-isometric deformations) datasets, using MATLAB on a Core-i7 system. Quantitative results, based on geodesic error curves, show the method achieves average geodesic errors of 0.071 on DT4D and 0.1718 on SMAL, outperforming baselines (GDD [48], IDPC [50], and Zoom Out [60]). Area Under geodesic error Curve (AUC) exceed 85% on DT4D and 62% on SMAL, with runtime around ~22 seconds per pair. Qualitative visualizations confirm superior matching in pose-variant and species-differing (non-isometric) shapes. The findings highlight enhanced accuracy and efficiency for non-rigid, non-isometric cases, attributed to intrinsic geodesic features and optimized assignment. Limitations include assumptions of complete meshes and potential sensitivity to extreme outliers; future work will extend to partial/degraded shapes and integrate post-processing for symmetry handling.

3. Conclusion

In this paper, an efficient method based on the LAPJV algorithm for linear assignment is proposed to establish unique correspondences between points of 3D shapes using geodesic distance features from salient points. Evaluations on benchmark datasets DT4D and SMAL demonstrate superior accuracy (lower average errors) compared to baselines (methods GDD [48], IDPC [50], and Zoom Out [60]), alongside reduced computational complexity and faster runtime. Overall, the approach excels in handling non-isometric deformations and pose variations.



دانشگاه صنعتی سهند

یک روش کارآمد برای تناظریابی اشکال سه بعدی با استفاده از ویژگی‌های

هندسی

DOR:

[20.1001.1.23223146.1404.12.1.1.4](https://doi.org/10.1001.1.23223146.1404.12.1.1.4)

نشریه علمی فناوری در مهندسی برق

دوره ۱۲ - شماره ۱

بهار و تابستان ۱۴۰۴

صفحات ۵ الی ۳۵

ISSN: 2322-3146

<http://jnsee.sut.ac.ir>

امیررضا امیرفتحیان^۱، حسین ابراهیم نژاد^{۲*}

^۱دانشجوی دکتری مهندسی برق- مخابرات، دانشکده مهندسی برق، دانشگاه صنعتی سهند تبریز، a_amirfathiyan@sut.ac.ir

^۲نویسنده مسئول، استاد گروه برق- مخابرات، دانشکده مهندسی برق، دانشگاه صنعتی سهند تبریز، ebrahimnezhad@sut.ac.ir

تاریخ پذیرش: ۱۴۰۳/۹/۷

تاریخ بازنگری: ۱۴۰۳/۷/۲۵

تاریخ ارسال: ۱۴۰۳/۳/۲۹

چکیده

در این مقاله یک روش کارآمد برای یافتن تناظر بین مدل‌های سه بعدی حاوی تغییر شکل‌های حرکتی و غیرصلب پیشنهاد شده است که از ویژگی‌های هندسی مدل‌ها استفاده می‌کند. برای این منظور، از یک ویژگی ذاتی مبتنی بر فاصله‌های ژئودزیک نقاط استفاده کرده و اثر بخشی روش پیشنهادی برای تناظریابی اشکال سه بعدی نشان داده می‌شود. ابتدا، مجموعه‌ای از نشانه‌ها (نقاط برجسته یا نقاط ویژه) از طریق نمونه برداری دورترین نقطه، شناسایی و انتخاب می‌شوند. سپس، ماتریس ویژگی مدل‌ها بر اساس فاصله ژئودزیک نقاط اشکال سه بعدی از نشانه‌ها ایجاد می‌شود. در ادامه، با به کارگیری الگوریتم LAPJV یک مسئله تخصیص خطی حل می‌شود که تابع هزینه آن حاوی ماتریس ویژگی مدل‌ها است. انجام آزمایش‌های گسترده روی مدل‌های مختلف از مجموعه داده‌های استاندارد نشان از برتری کمی و کیفی روش پیشنهادی در مقایسه با روش‌های دیگر از جمله رویکردهای [۴۸]، [۵۰] و [۶۰] دارد. نتایج نشان می‌دهند که روش پیشنهادی در مقایسه با روش‌های [۴۸]، [۵۰] و [۶۰] روی مدل‌های با تغییر شکل‌های حرکتی و غیرصلب و موارد غیرایزومتریک به نتایج بهتری در دقت و کارایی دست پیدا می‌کند.

واژه‌های کلیدی

تناظریابی اشکال سه بعدی،
آشکارساز نقاط برجسته،
توصیفگر فاصله ژئودزیک،
مسئله تخصیص خطی با الگوریتم
LAPJV،
تغییر شکل غیرایزومتریک.

۱- مقدمه

یکی از اساسی‌ترین کارهایی که در تحلیل اشکال سه‌بعدی در نظر گرفته می‌شود، یافتن نقاط متناظر بین دو یا چند شکل معین است. به عبارت دیگر، هدف اصلی مسئله تناظریابی پیدا کردن یک رابطه معنادار (نگاشت) بین عناصر اشکال داده شده است. تعیین اینکه چه تطابقی معنادار است به کار مورد نظر بستگی دارد. این کار می‌تواند از موارد ساده‌تر، شناسایی قسمت‌هایی از اشکال که از نظر هندسی مشابه است تا مسایل پیچیده‌تر، ارتباط بین عناصری از اشکال که بخش‌های یکسانی را نشان می‌دهند؛ متغیر باشد. البته ممکن است تکه‌های تطبیق به‌طور قابل توجهی از نظر هندسه، ساختار تکه در کل شکل یا توپولوژی متفاوت باشند. در شکل ۱ نمونه‌ای از چنین مسئله‌ای در تناظریابی نشان داده شده است. به‌طور کلی، محاسبه تناظرهای بین اشکال سه‌بعدی چالش‌برانگیز است، زیرا نیازمند درک عمیق ساختار اشکال و رفتار قطعات آن‌ها در مقیاس محلی و کلی (جهانی) است تا بتوان تناظرهای معناداری برقرار کرد. با وجود پیچیدگی ذاتی این مسئله، تناظریابی اشکال سه‌بعدی نقشی اساسی در کاربردهای متنوع پردازش هندسی ایفا می‌کند [۱]. از سوی دیگر، شناسایی خودکار و دقیق تناظرها بین دو یا چند شکل همچنان چالش مهمی است که بستری برای تحلیل‌های پیشرفته‌تر و کاربردهای گسترده‌تر در حوزه‌های مختلف را فراهم می‌کند. از این رو، تطبیق بین اشکال سه‌بعدی عملیاتی بنیادین در بسیاری از الگوریتم‌های گرافیک کامپیوتری و پردازش هندسی از جمله درونیابی اشکال^۱، بازسازی شکل^۲، انتقال بافت^۳، بازیابی شکل^۴، انتقال بخش‌ها^۵، هم‌ترازی داده‌های اسکن سه‌بعدی^۶، انتقال تغییرات شکل^۷، آشکارسازی تقارن^۸، آشکارسازی تغییرات^۹ و مدل‌سازی آماری^{۱۰} محسوب می‌شود [۲].



شکل ۱. یک مثال از مسائل چالشی در تناظریابی برای مجموعه‌ای از اشکال (ظروف دست‌ساز انسان). اجزای تشکیل دهنده اشکال از نظر هندسی و تعداد تکه‌های تشکیل دهنده (نوع و تعداد دستگیره، گردنه، تکیه‌گاه، دسته کوتاه یا بلند و نحوه اتصال به همدیگر) با هم متفاوت هستند [۱].

در اکثر روش‌های تناظریابی، توصیفگرهای پیشنهادی مبتنی بر توزیع هستند که از هیستوگرام برای بازنمایی ویژگی‌های محلی مدل‌های سه‌بعدی استفاده می‌کنند [۳]. این ویژگی‌ها شامل خصوصیات مهم اشکال هستند که در برابر تبدیل‌های هندسی نظیر مقیاس‌پذیری، چرخش، تاب خوردگی ناشی از جابه‌جایی اشکال، تغییرات روشنایی، مات‌شدگی و نویز مقاوم هستند. به‌طور کلی، فرآیند تناظریابی شامل سه مرحله اصلی است: استخراج ویژگی‌ها، استفاده از توصیفگر مناسب و برقراری تناظر بین ویژگی‌ها [۴].

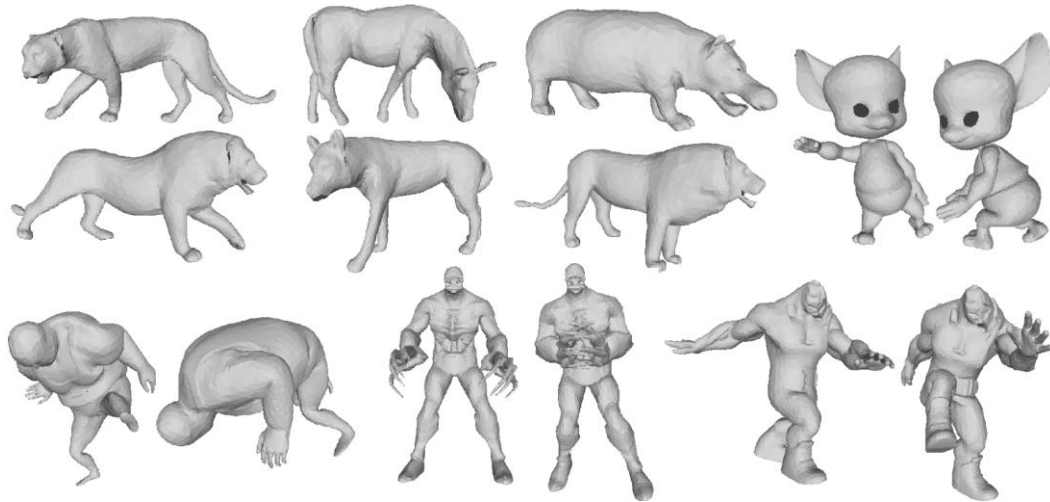
به‌طور سنتی مسئله تناظریابی در زمینه تجزیه و تحلیل تصاویر و حجم‌ها مورد مطالعه قرار گرفته‌است، اما امروزه در گرافیک کامپیوتری و برای اشکال هندسی با شبکه‌های مثلثی یا مجموعه نقاط، به‌عنوان یک هدف پژوهشی مطرح است. مسائل تناظریابی

¹ Shape interpolation
² Shape reconstruction
³ Texture transfer
⁴ Shape retrieval
⁵ Segmentation transfer
⁶ 3D scan alignment
⁷ Deformation transfer
⁸ Symmetry detection
⁹ Change detection
¹⁰ Statistical modeling

اشکال را می‌توان به انواع مختلفی طبقه‌بندی کرد. اما، تفاوت‌های قابل توجهی بین این مسائل وجود دارد که الزامات محاسباتی متفاوتی را برای حل آن‌ها ایجاد می‌کند. با وجود این تفاوت‌ها، همه این مسائل به هدف مشترکی، یعنی ایجاد نگاشتی معنادار از یک شکل به شکل دیگر، متمرکز هستند [۱]. به‌عنوان مثال، بر اساس مطالعات و بحث‌های ارائه شده در [۱] و روش معرفی شده در [۵]، مسئله بازیابی شکل ارتباط نزدیکی با تناظریابی اشکال دارد، زیرا می‌توان از تطابق عناصر (نظیر نقاط، وجه‌ها، نقاط ویژه و مشخصه‌های اسکلتی) برای سنجش شباهت بین دو شکل استفاده کرد.

روش‌های تناظریابی اشکال سه‌بعدی را می‌توان بر اساس معیارهای مختلفی از جمله سطح شباهت (کامل یا جزئی)، نوع پردازش شکل (جفتی یا مجموعه‌ای)، تراکم خروجی (پراکنده یا انبوه) و نوع تغییر شکل (ایزومتريک یا غیرایزومتريک) طبقه‌بندی کرد. با در نظر گرفتن این معیارها، فرآیند تطبیق بین اشکال سه‌بعدی به‌طور قابل توجهی پیچیده می‌شود. برای مثال، تناظریابی مبتنی بر تغییرات ایزومتريک ساده‌تر است، زیرا در این نوع تغییر شکل، فاصله‌های ژئودزیکی بین نقاط روی مش‌ها حفظ می‌شود. مفصل‌بندی، خمش و تبدیل‌های صلب نمونه‌هایی از تغییرات ایزومتريک هستند. در مقابل، تغییر شکل‌های غیرایزومتريک، مانند کشیدگی یا فشردگی، شباهت معنایی بین اشکال را حفظ می‌کنند، اما فواصل ژئودزیکی بین نقاط تغییر می‌یابد. اگرچه تناظریابی اشکال ایزومتريک به‌طور گسترده مورد مطالعه قرار گرفته است، ولی ممکن است استفاده از روش‌های مبتنی بر این سناریوها در موارد غیرایزومتريک عملکرد مطلوبی نداشته باشند. حتی با استفاده از روش‌های پیشرفته ارائه شده در [۶-۸]، دستیابی به عملکرد بهینه در سناریوهای غیرایزومتريک دشوار است. از سوی دیگر، انتخاب مجموعه داده‌ها نیز اهمیت بسزایی دارد، زیرا آن‌ها تحت تاثیر محدودیت‌ها و خطاهای مختلفی (نظیر نویز، خود انسدادی، همجوشی بین تکه‌ها) قرار دارند که به تکنیک‌های تصویر برداری از اشکال در دنیای واقعی و نوع روش‌های اسکن (تک منبعی یا چند منبعی) وابسته‌اند. این عوامل چالشی اضافی برای روش‌های تناظریابی اشکال ایجاد می‌کنند [۲]. یک چالش جدید زمانی پدید می‌آید که اشکال سه‌بعدی تحت تغییرات حرکتی شدید یا ترکیبی از حرکت (تغییر وضعیت) و تغییر شکل‌های غیرایزومتريک قرار گیرند. در چنین حالت‌هایی، هر چند ویژگی‌های غیرایزومتريک بین اشکال وجود دارد ولی فواصل ژئودزیکی در بخش‌هایی از آن‌ها به‌صورت تقریبی حفظ می‌شود. هدف اصلی این مطالعه، ارائه راه حلی مؤثر برای مواجهه با این چالش‌های پیچیده در تطبیق اشکال است. در شکل ۲، چالش‌های مختلف تطبیق بین اشکال سه‌بعدی به‌طور خلاصه نشان داده شده است.

در روش پیشنهادی، یک استراتژی کارآمد برای تطبیق مدل‌های سه‌بعدی بر اساس ویژگی‌های هندسی اشکال ورودی به کار گرفته شده است. روش پیشنهادی تنها نیازمند شناسایی اولین نقطه برجسته (نقطه عطف یا نقطه ویژه) مدل است. برای تعیین نقاط برجسته بعدی از نقاط برجسته قبلی استفاده می‌شود که منجر به کاهش بار محاسباتی روش می‌شود. با الهام از ساختار سیستم‌های موقعیت‌یابی ماهواره‌ای، ماتریس ویژگی‌های هر مدل به سادگی استخراج می‌شود.



شکل ۲. نمایشی از انواع چالش‌ها در تناظریابی اشکال سه‌بعدی.

در واقع، نقاط برجسته یا ویژه به عنوان نقاط مرجع در نظر گرفته می‌شوند و فاصله ژئودزیکی بین تمام جفت نقاط مدل و نقاط مرجع محاسبه می‌شود تا ماتریس ویژگی‌های مدل بر اساس توصیفگر ژئودزیکی به دست آید. برای دستیابی به یک رابطه صریح و تطبیق نهایی، استراتژی حل مسئله تخصیص خطی با الگوریتم LAPJV استفاده می‌شود که تابع هزینه آن اطلاعات ماتریس ویژگی مدل‌ها را در بر می‌گیرد. این استراتژی علاوه بر افزایش سرعت تطبیق، منجر به شناسایی منحصر به فرد هنگام تعیین نقاط هم‌مکان بین مدل‌ها می‌شود. زیرا، برخلاف الگوریتم‌های تکراری مانند k نزدیک‌ترین همسایه (KNN) یا نزدیک‌ترین نقاط تکراری (ICP) که ممکن است چندین نقطه متناظر در مدل هدف برای یک نقطه در مدل منبع پیدا کنند، روش پیشنهادی تضمین می‌کند که هر نقطه در مدل منبع مربوط به یک نقطه منحصر به فرد در مدل هدف است. روش پیشنهادی یک راه حل مؤثر برای چالش جدید در مسئله تطبیق شکل ارائه می‌دهد و عملکرد بهینه و پیشرفته‌ای را برای اشکال غیر صلب با تغییرات حرکتی شدید و موارد غیر ایزومتریک در مقایسه با سایر رویکردهای مشابه ارائه می‌دهد.

بقیه مقاله به شرح زیر سازماندهی شده است. در بخش ۲، برخی از مهمترین کارهای مربوط به تناظریابی بین اشکال سه‌بعدی بررسی شده است. روش پیشنهادی به تفصیل در بخش ۳ توضیح داده شده است. بخش ۴، به جزئیات پیاده‌سازی روش پیشنهادی اختصاص داده شده است. در بخش ۵، آزمایش و واکاوی نتایج روش پیشنهادی روی دو مجموعه داده استاندارد با خصوصیات مختلف به صورت کمی و کیفی در مقایسه با روش‌های پیشرفته مورد بحث قرار گرفته است. در نهایت، نتیجه‌گیری روش پیشنهادی در بخش ۶ ارائه شده است.

۲- کارهای مرتبط

تناظریابی اشکال سه‌بعدی یک حوزه تحقیقاتی شناخته شده‌ای است که سوابق موفقیت آمیز خوبی دارد و راهکارهای مختلفی برای حل چالش‌های آن اعم از تطبیق شکل ایزومتریک، غیر صلب یا غیر ایزومتریک ارائه شده است و مرور کلی تکنیک‌های مرتبط فراتر از محدوده این مقاله است. بنابراین، به بررسی و مطالعه آثاری تمرکز می‌شود که بیشترین ارتباط را با روش پیشنهادی دارند،

به‌ویژه روش‌هایی که بر تطبیق ویژگی‌های هندسی و توصیفگرهای هندسی مقاوم در برابر تغییرات شکل مدل‌ها استوار هستند. البته، خواننده علاقه‌مند می‌تواند برای مطالعه جامع‌تر در این زمینه به بررسی‌های اخیر در تناظریابی اشکال [۹، ۱۰، ۱۱] مراجعه نماید.

رویکردهای تطبیق ویژگی^۱: یکی از اساسی‌ترین روش‌ها برای تناظریابی بین اشکال، تخمین شباهت‌های بین عناصر یا ویژگی‌های خاصی از آنهاست که به عنوان رویکردهای تطبیق ویژگی یاد می‌شود. عناصر هر شکل اغلب توسط توصیفگرهای شکل تعریف می‌شوند و تناظرها با انتساب و بهینه‌سازی این توصیفگرها از طریق یک تابع هدف به دست می‌آیند. این تابع شامل به حداکثر رساندن شباهت بین توصیفگرهای مناظر و در عین حال به حداقل رساندن اعوجاج ناشی از تغییرات شکل است. در واقع، مسئله تناظریابی را می‌توان در قالب یک مسئله کمینه‌سازی بیان کرد که به طور موثری با استفاده از روش‌های تخصیص خطی قابل حل است. تطبیق ویژگی‌ها را می‌توان در زمینه‌های مختلفی که امکان محاسبه مجموعه‌ای از توصیفگرها برای عناصر اشکال وجود دارد، اعمال کرد و می‌تواند به عنوان مبنایی برای ایجاد تطابق اولیه در سایر رویکردها عمل کند. از جمله کاربردهای آن می‌توان به ثبت اسکن‌های سه‌بعدی، تغییر شکل سطوح و تطبیق اسکلتی اشاره کرد. علاوه بر این، استفاده از مدل‌های پراکنده^۲ برای تطبیق اشکال جزئی ایزومتریک در سطوح مترام^۳ [۱۱] رویکردی است که با تعداد محدودی از ویژگی‌های هندسی می‌خواهد به بهبود کارایی تطابق در سناریوهای پیچیده شامل تغییر شکل‌های محلی یا تغییرات هندسی، کمک کند. تجزیه و تحلیل مبتنی بر ویژگی برای بسیاری از وظایف پردازش شکل با نقاط ویژه که توسط توصیفگرهای محلی شناسایی می‌شوند، یک امر اساسی است. این توصیفگرها اغلب نواحی مورد علاقه را در اطراف نقاط منفرد تعریف می‌کنند. اما، استفاده از این نواحی می‌تواند تطابق جزئی را پیچیده کند. زیرا ممکن است به خوبی با هندسه قطعات شکل هماهنگ نباشند. روش ارائه‌شده در [۱۲] تکنیکی برای نگاشت دوطرفه معرفی می‌کند که در آن توصیفگر ناحیه مورد نظر با استفاده از دو نقطه ویژه تعریف می‌شود. این روش با تطبیق شکل ناحیه هدف، تناظریابی جزئی را تسهیل می‌کند و با نادیده گرفتن نواحی غیر ضروری بر اساس ویژگی‌های تطبیقی نگاشت دوطرفه، نگاشت حاصل را تا حدی در برابر تغییرات توپولوژیکی مقاوم می‌کند. تطبیق ویژگی‌ها اغلب به نشانه‌ها، نقاط کلیدی^۴ یا ویژگی‌های متمایزی از سطح شی متکی است که به عنوان نقاط مرجع برای مقایسه و تراز کردن اشکال سه‌بعدی به کار می‌روند. با شناسایی و تطبیق این نقاط بین دو یا چند شکل سه‌بعدی می‌توان شباهت‌ها، تفاوت‌ها و روابط بین آنها را تجزیه و تحلیل کرد که به رویکردهای مبتنی بر تطبیق ویژگی و تناظریابی پراکنده کمک می‌کند. در رویکردهای مبتنی بر نگاشت تابعی، نقاط کلیدی به‌عنوان قیدهای ویژه برای رفع ابهام تقارن استفاده می‌شوند [۱۴، ۱۳]. روش‌های متعددی از نشانه‌های نقطه‌ای استفاده می‌کنند که ممکن است به‌صورت دستی توسط کاربران تعریف شوند [۱۵] یا به‌طور خودکار محاسبه گردند [۱۶] یا حتی به محدودیت‌های منحنی [۱۷] گسترش یابند. این تکنیک‌ها با اتکا بر استراتژی‌های حفظ توصیفگرهایی مانند هسته حرارتی^۵ [۱۸] یا نگاشت‌های هسته موج^۶ [۱۹] و تمرکز بر کمینه‌سازی مربعات خطا در طول بهینه‌سازی به تطبیق نشانه‌ها دست می‌یابند. روش‌های تطبیق مبتنی بر نقاط کلیدی در سایر چارچوب‌های تناظریابی از جمله روش‌های مبتنی بر هم‌ترازی شکل بیرونی تحت قیدهای معین [۲۰-۲۲]، جاسازی اشکال در یک حوزه پارامتری مشترک [۲۳-۲۷]، پارامترسازی متقابل سطوح [۲۸]، کمینه‌سازی اعوجاج [۲۹] یا بهینه‌سازی مستقیم انرژی متقارن دیریکله [۳۰] برای حفظ ویژگی‌های

¹ Feature Matching Approaches

² Sparse

³ Dense

⁴ Landmark or Key points

⁵ Heat Kernel Signature

⁶ Wave Kernel Maps

متمايز (نشانه‌ها) نیز مورد بررسی قرار گرفته است. تکنیک‌های اخیر می‌توانند به نتایج دقیقی در تناظریابی منجر شوند. اما، اغلب از نظر محاسباتی سنگین بوده و قیدهای توپولوژیکی را در فرآیند تناظریابی تحمیل می‌کنند. یکی از جنبه‌های مهم رویکردهای مبتنی بر ویژگی، انتخاب ویژگی‌های مناسب است. بسیاری از مطالعات نظیر [۳۱] از توابع ویژه عملگر لاپلاس بلترامی^۱ استفاده می‌کنند که برای نمایش توابع صاف با تغییرات محدود بهینه هستند [۳۲]. با این حال، ماهیت کلی پایه‌های لاپلاس بلترامی کل شکل را در تحلیل ویژگی‌ها در نظر می‌گیرد و ممکن است به‌طور کامل با تغییرات شکل غیرایزومتريک سازگار نباشد. برای رفع این محدودیت، حالت‌های مینفولد فشرده [۳۳] معرفی شدند تا پراکندگی و محلی بودن را در ساخت پایه‌ها افزایش دهند. این رویکرد، به جای در نظر گرفتن کل شکل به صورت یکپارچه بر شناسایی و استفاده از اطلاعات محلی و مرتبط تمرکز دارد و استخراج ویژگی‌های شکل را بهبود می‌بخشد. با ترویج پراکندگی (وجود تعداد محدودی از عناصر غیرصفر) و محلی بودن (تمرکز بر نقاط نزدیک)، پایه‌های لاپلاس بلترامی می‌توانند به‌طور موثرتری برای ثبت تغییرات غیرایزومتريک تنظیم شوند. به همین منظور، عملگر همیلتونی^۲ [۳۴] برای استخراج ویژگی‌های محلی موثر معرفی شد. همچنین، روش‌هایی با ترکیب پایه‌های لاپلاس بلترامی و ویژگی‌های هندسی ارائه شده‌اند که انعطاف‌پذیری بیشتری فراهم می‌کنند [۳۵-۳۷] و نگاهی به تناظریابی را بهبود می‌بخشند. اما، این روش‌ها اغلب برای اشکال تقریباً ایزومتريک طراحی شده‌اند و فقط محدودیت‌های مربوطه را برآورده می‌کنند. توصیف قدرتمندی از هندسه اشکال از طریق هم‌ترازی مقادیر ویژه لاپلاسی بین اشکال غیرایزومتريک در [۳۸] معرفی شده است. در واقع، هم‌ترازی طیفی به صورت ذاتی دو شکل را ایزومتريک می‌کند و از این رو، می‌تواند تطبیق دقیق را با استفاده از روش‌های ایزومتريک موجود [۴۰، ۳۹] تسهیل کند. رویکردهای تطبیق توصیفگرها: روش دیگری برای تناظریابی، استفاده از شباهت بین توصیفگرهای عناصر اشکال است که می‌تواند از طریق تطبیق نقطه‌ای دستی [۴۱، ۴۲] یا خودکار [۴۳، ۴۴] حاصل شود. تطبیق بهینه ممکن است از توصیفگرهای نقطه به نقطه مانند توصیفگرهای فاصله [۴۵، ۴۶] یا توصیفگرهای شات^۳ [۴۷] استفاده کند. در روش‌های مبتنی بر شباهت، متغیرهای هندسی یا توصیفگرها می‌توانند تحت یک مدل تغییر شکل یافته محاسبه شوند که به صورت عبارات نقطه‌ای (روی رئوس) یا عبارات زوجی (بین جفت رئوس) تعریف می‌شوند. عبارات زوجی در این میان مؤثرتر هستند. زیرا، ترم‌های نقطه‌ای به محدودیت‌ها و ویژگی‌های فردی نقاط اشاره دارند و شباهت یا عدم شباهت بین نقاط متناظر را نشان می‌دهند. در حالی که ترم‌های زوجی روابط بین جفت نقاط را رمزگذاری کرده و اطلاعات مکانی یا ساختاری را ثبت می‌کنند. ترکیب این توصیفگرها به ایجاد یک تابع انرژی منجر می‌شود که با کمینه‌سازی آن، تطبیق مورد نظر را حاصل می‌شود و رویکردی دقیق‌تر و مقاوم‌تر برای حل مسائل تناظریابی فراهم می‌کند، در نهایت منجر به بهبود نتایج می‌شود [۱۰]. انتخاب توصیفگر مناسب به نوع تغییر شکل بستگی دارد. برای مثال در تغییرات ایزومتريک، استفاده از فواصل ژئودزیکی مناسب است. زیرا این فواصل بین جفت نقاط روی سطح اشکال ایزومتريک حفظ شده و فرآیند تطبیق را تسهیل می‌کنند. به عنوان نمونه در مطالعه [۴۸]، GDD^۴ یک توصیفگر فاصله ژئودزیکی معرفی شده است که فواصل ژئودزیکی را به صورت ترکیبی خطی از توابع پایه کدگذاری می‌کند. این رویکرد یک پایه فاصله ژئودزیکی را تعریف می‌کند که برای تقریب فشرده فاصله‌های ژئودزیکی با استفاده از نرم فروبینوس بهینه‌سازی شده و نمایش دقیق‌تری ارائه می‌دهد. در واقع، تناظریابی را می‌توان

¹ Laplace-Beltrami Operator

² Hamiltonian Operator

³ Signature of Histograms of Orientations (SHOT)

⁴ Geodesic Distance Descriptors

با مقایسه توصیفگرهای فاصله ژئودزیکی اشکال از طریق روش‌های جستجوی نزدیکترین همسایه مانند الگوریتم ICP^۱، بدون نیاز به روش‌های برش یا سست‌سازی انجام داد. این عملیات نشان‌دهنده رویکردی مستقیم برای تطبیق اشکال بر اساس نمایش‌های GDD آن‌ها است.

روش‌های نگاشت‌های تابعی^۲: نگاشت‌های تابعی معرفی شده در [۴۹]، مسئله تناظریابی را به صورت نگاشتی بین توابع تعریف شده روی سطوح، به جای نگاشت مستقیم نقاط سطح، مدل‌سازی می‌کنند که رویکردی بسیار مؤثر است. مزیت اصلی این روش، کاهش قابل توجه ابعاد مسئله تطبیق است، زیرا از پایه‌های ویژه عملگر لاپلاس بلترامی برای نمایش این توابع استفاده می‌شود. با گذشت زمان، مفهوم نگاشت‌های تابعی برای کاربردهای مختلف بهبود یافته [۵۰-۵۳] و به تازگی در قالب یادگیری عمیق ادغام شده است [۵۵، ۵۴]. برخلاف رویکردهای سنتی مبتنی بر توصیفگر، این روش‌ها بر بهینه‌سازی توابعی تمرکز دارند که در فضایی با ابعاد پایین‌تر نسبت به تعداد رئوس (نقاط) اشکال، تجزیه و تحلیل و نمایش داده می‌شوند. پس از محاسبه نگاشتهای تابعی، تناظرهای بین رئوس با استفاده از روش‌های مختلف [۴۹] استخراج می‌شوند که با برجسته‌سازی شباهت‌ها و تفاوت‌ها در هندسه اشکال، مقایسه دقیق‌تری را امکان‌پذیر می‌کنند. یک استراتژی رایج برای بهبود عملکرد تناظریابی، اصلاح نگاشت‌های تابعی است که اغلب از تکنیک ICP برای تبدیل نگاشت‌های تابعی به نقطه‌ای استفاده می‌کند. در ادامه، بهینه‌سازی نگاشت‌ها از طریق کمیته‌سازی یک تابع انرژی [۵۶-۵۸]، انجام تناظریابی مؤثر با شرط دوطرفه بودن نگاشت به عنوان یک الزام اساسی [۵۹]، افزایش وضوح طیفی نگاشت‌های تابعی در طول پالایش [۶۰] و تجزیه و تحلیل رابطه بین نگاشت‌های تابعی و نقطه‌ای با ارائه روش اصلاحی مبتنی بر هم‌ترازی طیفی [۶۱] معرفی شده‌اند. با این حال، روش‌های دو مرحله‌ای تطبیق و پالایش به دلیل احتمال خطا در تعیین نگاشت‌های تابعی اولیه، اغلب عملکرد کمتر از حد مطلوب دارند. در سال‌های اخیر، محققان به‌طور فزاینده‌ای بر روش‌های یادگیری عمیق برای نگاشت‌های تابعی تمرکز کرده‌اند. از سال ۲۰۱۷ مطالعات متعددی نگاشت‌های تابعی عمیق را معرفی کرده‌اند، از جمله شبکه‌های عمیق نظارت شده با نگاشت تابعی موسوم به FMNet^۳ [۶۲]، نسخه‌های بدون نظارت FMNet [۶۳] و شبکه‌های یادگیری عمیق DUO-FMNet^۴ [۶۴] که هدف اصلی این رویکردها، یادگیری مستقیم ویژگی‌ها از داده‌های آموزشی است.

۳- روش پیشنهادی

با وجود پیشرفت‌های قابل توجه در زمینه تناظریابی اشکال، تطبیق بین مدل‌های غیرایزومتریک همچنان چالشی اساسی است. مطالعات نشان می‌دهد که دشوارترین حالت در تطبیق اشکال سه بعدی زمانی رخ می‌دهد که اشکال از نظر معنایی به یک کلاس تعلق داشته باشند، اما لازم نیست که ایزومتریک باشند. در چنین مواردی، الگوریتم تناظریابی باید قادر به شناسایی و ثبت ویژگی‌های معنادار در هر دو شکل باشد و همزمان اعوجاج محلی را کاهش دهد تا نتایجی با کیفیت بالا ارائه کند. بررسی‌ها حاکی از آن است که با وجود پیشرفت سریع روش‌های نگاشت‌های تابعی مبتنی بر یادگیری عمیق، رویکردهای بدون نظارت موجود بیشتر بر تطبیق اشکال نزدیک به ایزومتریک تمرکز دارند یا در مقایسه با روش‌های نظارت شده در تطبیق اشکال غیرایزومتریک با شکاف‌های

^۱ Iterative Closest Point Algorithm

^۲ Functional Map Methods

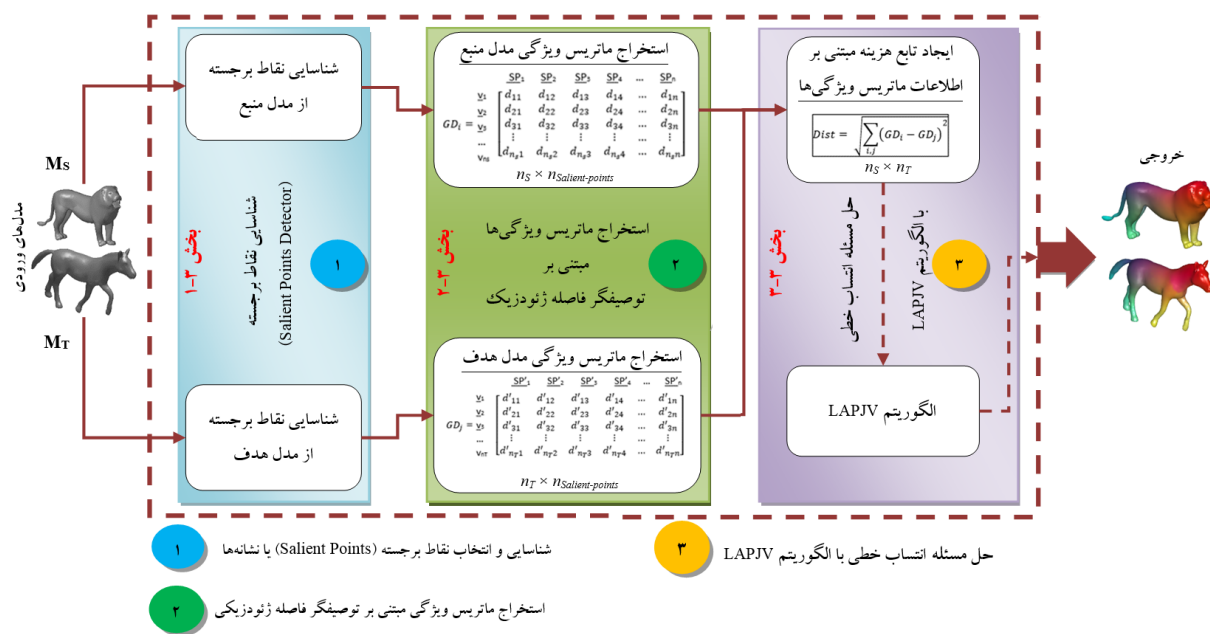
^۳ Functional Map Networks

^۴ Deep Orientation-aware Functional Map Networks

عملکردی قابل توجهی مواجه هستند [۶۵]. در این مقاله، تلاش شده است تا راه‌حلی موثر برای تطبیق اشکال سه‌بعدی با تغییرات حرکتی شدید (غیر صُلب) یا مدل‌هایی که به‌طور همزمان شامل تغییر موقعیت و تغییر شکل غیرایزومتریک هستند، ارائه شود. رویکرد پیشنهادی ارتباط نزدیکی با روش‌های مبتنی بر تطبیق ویژگی‌ها و توصیفگرها دارد و به‌طور مستقیم نداشت‌های نقطه‌ای را تولید می‌کند که تا حدی بهینه بوده و بر حفظ نقاط برجسته در حین محاسبه نداشت کلی تاکید دارد. هدف، دستیابی به تطبیق نقطه‌ای با کیفیت بالا بین دو شکل سه‌بعدی، یعنی شکل منبع (M_s) و شکل هدف (M_T)، است که به‌صورت مش‌های مثلثی چند وجهی (منیفولد) با تعداد رئوس n_s و n_T (تعداد نقاط معلوم) در دسترس هستند. روش پیشنهادی شامل سه بخش اصلی است:

- (۱) شناسایی نقاط برجسته (نقاط ویژه) اشکال ورودی
- (۲) استخراج ماتریس ویژگی مدل‌ها: بر اساس توصیفگر فاصله ژئودزیک و نقاط برجسته استخراج شده
- (۳) استراتژی حل مسئله تخصیص خطی با الگوریتم LAPJV^۱ برای تطبیق نهایی

روش پیشنهادی روی مجموعه داده‌های پنج مارک متنوع غیر صُلب و غیرایزومتریک ارزیابی شده و نتایج آن با رویکردهای مشابه مبتنی بر تطبیق ویژگی با توصیفگر ژئودزیک و نداشت‌های تابعی مقایسه شده است. روش پیشنهادی بهبود قابل توجهی در تطبیق اشکال سه‌بعدی غیر صُلب و غیرایزومتریک نسبت به روش‌های رقابتی (که در ادامه مقایسه خواهند شد) ایجاد کرده و به کاهش پیچیدگی محاسباتی، افزایش سرعت و بهبود کارایی تطبیق منجر شده است. بلوک دیاگرام و طرح کلی روش پیشنهادی، شامل تعاملات و روابط بین اجزای مختلف در شکل ۳ ارائه شده است. در ادامه، جزئیات کامل اجرای هر یک از مراحل روش پیشنهادی به‌طور جامع شرح داده شده است.



شکل ۳. بلوک دیاگرام و طرح کلی روش پیشنهادی برای تناظریابی بین اشکال سه‌بعدی.

¹ Linear Assignment Problem solver with Junker Volgenant algorithm (LAPJV)

۱-۳ - شناسایی نقاط برجسته (نقاط ویژه) اشکال ورودی

بر اساس تحلیل ادراک انسانی در حل مسئله تطبیق، می‌توان نقاطی با ویژگی‌های برجسته و متناظر در هر مدل سه‌بعدی شناسایی کرد که به صورت شهودی قابل فهم هستند. این نقاط که در نواحی با تغییرات انحنا قوی و پایداری محلی قرار دارند، اغلب روی تفرع‌ها و تحدب‌های اشکال واقع شده‌اند [۶۶]. دقت وجوه مثلثی در این نقاط باید بیشتر باشد، به همین دلیل این نواحی از مثلث‌های کوچک‌تر تشکیل می‌شوند. این نقاط به‌عنوان ماکزیمم‌های محلی یک تابع تعریف شده روی سطوح عمل می‌کنند که ویژگی‌های مختلف شکل را توصیف کرده و تغییرات محلی آن را نشان می‌دهد که برای ادراک انسانی اهمیت ویژه‌ای دارد. یکی از ویژگی‌های این نقاط، فاصله زیاد آن‌ها از یکدیگر است، به طوری که نسبت به سایر نقاط مدل، بیشترین فاصله ژئودزیکی را دارند. با توجه به این خصوصیات، این نقاط به‌عنوان نقاط برجسته یا ویژه شناخته می‌شوند. استفاده از نقاط برجسته در توصیف ویژگی‌های مدل‌های سه‌بعدی برای الگوریتم‌های تطبیق شکل مفید است، زیرا این نقاط نمایش فشرده‌ای ارائه داده و نسبت به تبدیل‌های هندسی مقاوم هستند. روش پیشنهادی با الهام از ساختار سیستم‌های موقعیت‌یابی ماهواره‌ای و استفاده از این نقاط به‌عنوان نقاط مرجع، قادر به تعیین دقیق نقاط هم‌مکان بین مدل‌ها است. نقاط هم‌مکان، نقاطی روی مدل هستند که فاصله یکسانی از نقاط برجسته مربوطه دارند. در روش پیشنهادی، زیر مجموعه‌ای از نقاط هر مدل با استفاده از استراتژی نمونه‌برداری دورترین نقطه مبتنی بر فاصله ژئودزیکی شناسایی می‌شوند. این روش با بهره‌گیری از الگوریتم Fast Marching [۶۷]، فواصل ژئودزیکی نقاط را محاسبه کرده و تعدادی از آن‌ها را که در قسمت‌های مختلف مدل دارای بیشترین فاصله ژئودزیکی هستند، انتخاب می‌کند. اما، محاسبه فاصله ژئودزیکی بین تمام جفت نقاط به افزایش پیچیدگی محاسباتی منجر می‌شود. برای رفع این مشکل، می‌توان میانگین فاصله ژئودزیکی، AGD^1 ، بین تعداد محدودی از نقاط مدل را محاسبه و از آن برای شناسایی نقاط برجسته استفاده کرد. در روش پیشنهادی، معیار انتخاب نقاط ویژه به صورت تکراری و بر اساس استراتژی نمونه‌برداری دورترین نقطه از سایر نقاط مش انجام می‌شود. این فرآیند با انتخاب یک نقطه تصادفی روی مدل آغاز شده و تا شناسایی p_s نقطه ویژه ادامه می‌یابد. ابتدا، میانگین فاصله‌های ژئودزیکی بین نقاط مش از یک نقطه تصادفی محاسبه شده و نقطه با بیشترین مقدار فاصله به‌عنوان اولین نقطه برجسته (نقطه ویژه) تعیین می‌شود. برای شناسایی دومین نقطه برجسته، فواصل ژئودزیکی بین اولین نقطه ویژه و سایر نقاط مدل محاسبه شده و نقطه‌ای با بیشترین فاصله ژئودزیکی به‌عنوان نقطه برجسته دوم انتخاب می‌شود. الگوریتم پیشنهادی به همین ترتیب تکرار می‌شود تا نقاط ویژه بعدی استخراج شوند. در واقع، روش پیشنهادی از مجموعه نقاط برجسته قبلی برای یافتن نقطه ویژه بعدی استفاده می‌کند و این فرآیند تا رسیدن به تعداد مورد نظر نقاط ویژه، p_s ادامه می‌یابد. این رویکرد با هدف ارائه تقریب دقیقی از میانگین فاصله ژئودزیکی با پیچیدگی محاسباتی کمتر طراحی شده است به طوری که نیازی به محاسبه فواصل بین تمام جفت نقاط مدل ندارد. انتخاب مقدار p_s از اهمیت ویژه‌ای برخوردار است، زیرا مقادیر بزرگ p_s تعداد نقاط مورد استفاده در تقریب را افزایش داده و در نتیجه پیچیدگی محاسباتی را بالا می‌برد. در مقابل، انتخاب مقادیر کوچک برای p_s ، تعداد نقاط محدودی را در فرآیند تقریب دخیل می‌کند که ممکن است به دلیل کمبود اطلاعات کافی درباره شکل، منجر به تخمین نادرست فاصله شود. در روش پیشنهادی با توجه به ساختار توپولوژیکی اشکال ورودی و نتایج آزمایش‌ها، مقدار p_s برابر با ۱۶ انتخاب شده است. بر اساس فرض اولیه، هر یک از اشکال سه‌بعدی به صورت یک منیفولد در نظر

¹ Average Geodesic Distance

گرفته می‌شود. یک منیفولد یا یک شبکه چند وجهی مثلثی به‌عنوان مجموعه‌ای از رئوس، وجوه و یال‌های همسایگی مرتبه اول رئوس یا لبه‌ها تعریف شده و با نماد $M = \{V, F, E\}$ نشان داده می‌شود. این المان‌ها مطابق با معادله (۱) تعریف می‌شوند.

$$\begin{aligned} V &= \{v_i \mid v_i \in \mathcal{R}^3, 1 < i \leq n\} \\ F &= \{f_{ijk} = (v_i, v_j, v_k) \mid v_i, v_j, v_k \in V, i \neq j \neq k\} \\ E &= \{e_{ij} = (v_i, v_j) \mid v_i, v_j \in V, i \neq j\} \end{aligned} \quad (1)$$

در رابطه (۱)، V مجموعه رئوس، F مجموعه وجوه و E مجموعه یال‌های همسایگی مرتبه اول رئوس برای یک مش با n رأس را نشان می‌دهند. برای شناسایی نقاط برجسته از میانگین فاصله ژئودزیکی، AGD، استفاده می‌شود که تقریبی از آن طبق رابطه (۲) محاسبه می‌شود.

$$AGD = f(v) \approx \frac{1}{n_{ps}} \sum_{i=1}^{n_{ps}} D_{Geodesic}(v, p_i) \quad (2)$$

در رابطه (۲)، n_{ps} تعداد نقاط تصادفی، v رئوس و p_i هر نقطه تصادفی روی مش مثلثی را نشان می‌دهد. اولین نقطه برجسته مدل با تعیین حداکثر مقدار تابع میانگین فاصله ژئودزیکی AGD طبق رابطه (۳) تعیین می‌شود.

$$sp_1 = \underset{v \in V}{\text{Argmax}}(AGD(v)) \quad (3)$$

در رابطه (۳)، sp_1 اولین نقطه برجسته مدل را نشان می‌دهد که در واقع اندیس رأسی از مدل را مشخص می‌کند. در الگوریتم پیشنهادی، پس از شناسایی اولین نقطه برجسته، استخراج نقاط ویژه بعدی به سادگی انجام می‌شود. به‌طور کلی، استراتژی استفاده از نقاط ویژه قبلی برای یافتن نقطه ویژه بعدی در نظر گرفته می‌شود. برای شناسایی نقاط برجسته (ویژه) هر مدل از رابطه (۴) استفاده می‌شود.

$$sp_m = \underset{v \in V}{\text{Argmax}}(\min(D_{Geodesic}(v, sp_1), D_{Geodesic}(v, sp_2), \dots, D_{Geodesic}(v, sp_{m-1}))) \quad (4)$$

در رابطه (۴)، $sp_1, sp_2, sp_3, \dots, sp_{m-1}$ نقاط برجسته (ویژه) شناسایی شده و v رئوس مش مثلثی هستند که برای شناسایی نقطه ویژه جدید، sp_m ، استفاده می‌شوند. در واقع، فواصل ژئودزیکی هر راس مش (v) از نقاط ویژه شناسایی شده (sp_{m-1}) محاسبه شده و کمینه این فاصله‌ها انتخاب می‌شود. سپس از میان نقاط با کمینه فواصل، نقطه‌ای که دارای بیشترین فاصله ژئودزیکی است به‌عنوان نقطه برجسته (ویژه) جدید انتخاب می‌شود. استفاده از الگوریتم نمونه‌برداری دورترین نقطه، احتمال انتخاب نقاط نزدیک به نقاط برجسته قبلی را کاهش می‌دهد، زیرا این نقاط به دلیل همسایگی با نقاط برجسته قبلی فواصل کمتری دارند. این رویکرد از انتخاب چندین نقطه برجسته در یک ناحیه از مدل جلوگیری کرده و تضمین می‌کند که نقاط انتخاب شده، دورترین نقاط نسبت به یکدیگر باشند. در نتیجه نقاط برجسته در نواحی متمایز و برجسته‌تر مدل قرار می‌گیرند. الگوریتم مورد استفاده برای استخراج نقاط برجسته (نقاط ویژه) در الگوریتم ۱ خلاصه شده است. پس از شناسایی این نقاط، یک تطبیق اولیه پراکنده بین آن‌ها برقرار می‌شود. برای این منظور، می‌توان از روش‌های دستی یا خودکار مانند توصیفگرهای HKS [۱۸] یا WKS [۱۹] استفاده کرد. از آنجا که تطبیق با این ابزارها می‌تواند زمان‌بر بوده و طول مدت تطابق متراکم را طولانی‌تر کند؛ تطبیق اولیه به‌صورت دستی و بر اساس شهود انسانی و ادراک بصری برقرار می‌شود.

الگوریتم ۱: الگوریتم استفاده شده برای شناسایی نقاط برجسته از مدل ورودی

ورودی‌ها: - مدل‌های ورودی مرجع و هدف. (مش‌های مثلثی چندوجهی مجموعه‌ای از رأس‌ها، وجه‌ها و یال‌ها)

- تعداد نقاط برجسته مورد نیاز (N_{sp})

خروجی: بردار نقاط ویژه مدل حاوی اندیس نقاط برجسته (F_p)

۱: انتخاب n نقطه تصادفی از مش ورودی. (در اینجا $n=16$)

۲: **حلقه for** با شمارنده ۱ تا تطبیق

۳: محاسبه فاصله ژئودزیک هر نقطه از همه نقاط روی مش با الگوریتم پیمایش سریع (Fast-Marching)

۴: **پایان حلقه for**

۵: محاسبه میانگین فواصل ژئودزیک $(AGD \leftarrow \frac{1}{n} \sum_{i=1}^n D_{Geodesic}(v, p_i))$

۶: شناسایی نقطه با بیشترین مقدار میانگین فاصله ژئودزیک به عنوان اولین نقطه برجسته $(sp_1 \leftarrow \text{Argmax}_{v \in V}(AGD(v)))$

۷: شناسایی دومین نقطه برجسته مبتنی بر نقطه برجسته اول $(sp_2 = \text{Argmax}_{v \in V}(D_{Geodesic}(v, sp_1)))$

۸: **حلقه for** با شمارنده ۳ تا N_{sp}

۹: شناسایی نقاط بعدی مبتنی بر نقاط برجسته قبلی

$(sp_m = \text{Argmax}_{v \in V}(D_{Geodesic}(v, sp_1), D_{Geodesic}(v, sp_2), \dots, D_{Geodesic}(v, sp_{m-1})))$

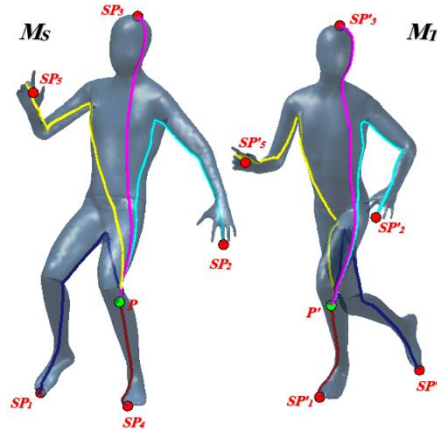
۱۰: **پایان حلقه for**

۱۱: استخراج اندیس نقاط برجسته به صورت یک بردار ویژگی (F_p)

۳-۲- استخراج ماتریس ویژگی مدل‌ها

روش پیشنهادی برای استخراج ویژگی‌های هر مدل از نقاط برجسته (نقاط ویژه) شناسایی شده استفاده می‌کند. این فرآیند با الهام از ساختار سیستم‌های موقعیت‌یابی جهانی، GPS، بر اساس محاسبه فواصل ژئودزیک تمام نقاط مدل از نقاط ویژه مربوطه انجام می‌شود. برای تعیین موقعیت هر نقطه مدل، فاصله ژئودزیک آن نقطه تا نقاط برجسته با استفاده از الگوریتم Fast Marching محاسبه می‌شود. این الگوریتم برای محاسبه کوتاه‌ترین مسیر یا فاصله ژئودزیک روی سطوح مختلف از جمله مش‌های مثلثی به کار می‌رود و به عنوان روشی برای محاسبه فاصله‌های ژئودزیک ریمانی عمل می‌کند، جایی که کوتاه‌ترین مسیر به صورت منحنی روی سطح نمایش داده می‌شود. بر این اساس، موقعیت هر نقطه فرضی p در مدل به صورت زیر تعیین می‌شود: مکان نقطه p نسبت به نقطه برجسته sp_1 ، محیط کره‌ای به مرکز sp_1 و شعاع فاصله ژئودزیک از نقطه p تا sp_1 است. به همین ترتیب، مکان نقطه p نسبت به نقطه برجسته sp_2 ، محیط کره‌ای به مرکز sp_2 و به شعاع فاصله ژئودزیک از نقطه p تا sp_2 تعیین می‌شود. با جمع‌آوری این اطلاعات، موقعیت نسبی نقطه p روی مدل مشخص می‌شود. به طور کلی، مکان جدید از تلاقی این کره‌ها به دست می‌آید که محل تقاطع آن‌ها دایره‌ای را تشکیل می‌دهد و نقطه p روی محیط این دایره قرار می‌گیرد. این فرآیند برای سایر نقاط برجسته نیز تکرار می‌شود. از نظر تئوری، موقعیت نقطه p با استفاده از سه نقطه برجسته (ویژه) قابل محاسبه است، اما برای دستیابی به دقت بالاتر، حداقل از چهار نقطه برجسته استفاده می‌شود. با افزایش تعداد نقاط برجسته، تقاطع کره‌ها به یک نقطه واحد همگرا می‌شود که مختصات مکانی دقیق آن نقطه را نشان می‌دهد. فواصل ژئودزیک بر اساس نگاشت $d(v): V \rightarrow R$ محاسبه می‌شوند. در این مرحله، ماتریس ویژگی بر اساس

فواصل ژئودزیک هر نقطه از نقاط برجسته با استفاده از الگوریتم ۲ محاسبه می‌شود. در شکل ۴ خروجی الگوریتم ۲ و نمایشی از برقراری تطبیق بین نقاط فرضی در مدل‌ها نشان داده شده است. به عبارت دیگر، شرط تطبیق نقطه فرضی p در مدل منبع با نقطه p' در مدل هدف این است که فواصل نقطه p از نقاط ویژه مدل منبع $(sp_1, sp_2, sp_3, \dots, sp_m)$ برابر با فواصل نقطه p' از نقاط ویژه متناظر در مدل هدف $(sp'_1, sp'_2, sp'_3, \dots, sp'_m)$ باشد. ابعاد ماتریس ویژگی هر شکل که با نماد **GD** نشان داده می‌شود، برابر با تعداد نقاط (رئوس) در تعداد نقاط ویژه هر شکل است. الگوریتم ۲ خلاصه‌ای از روش به کار رفته را ارائه می‌دهد.



شکل ۴. نمایشی از برقراری تناظر بین نقاط مدل‌ها بر اساس فاصله ژئودزیک هر نقطه (مانند p و p') از نقاط ویژه مدل‌ها ($m=5$).

الگوریتم ۲: استخراج ماتریس ویژگی مدل‌های ورودی مبتنی بر فاصله‌های ژئودزیک.

ورودی‌ها: - مدل‌های ورودی مرجع و هدف. (مش‌های مثلثی چندوجهی مجموعه‌ای از رأس‌ها، وجه‌ها و یال‌ها).

- زیرمجموعه‌ای از رئوس مش‌های ورودی $S \subset V$

- در اینجا از نقاط برجسته (نقاط ویژه) تعیین شده استفاده می‌شود.

خروجی: ماتریس ویژگی مدل (ماتریس فاصله ژئودزیک نرمالیزه، **GD**)

مقداردهی اولیه داده‌ها

۱: تنظیم نقاط برجسته به عنوان نقاط شروع برای محاسبه فاصله‌های ژئودزیک $(pstarts \leftarrow \{sp_i\}_{i=1}^{N_{sp}})$

۲: انتخاب تنظیم تصادفی نقطه انتهایی (se) برای ترسیم مسیرهای ژئودزیک $(N_{se} \leftarrow 2)$

بدنه اصلی

۳: **حلقه for** با شمارنده ۱ تا $\{sp_i\}_{i=1}^{N_{sp}}$

۴: محاسبه فاصله‌های ژئودزیک با الگوریتم انتشار پیشرو Fast Marching

۵: نرمالیزه کردن فاصله‌های ژئودزیک محاسبه شده

۶: **پایان حلقه for**

۷: استخراج ماتریس ویژگی هر مدل (ماتریس فاصله ژئودزیک نرمالیزه، **GD**)

۳-۳ - استراتژی حل مسئله تخصیص خطی با الگوریتم LAPJV برای تطبیق نهایی

در مرحله نهایی روش پیشنهادی از استراتژی حل مسئله تخصیص خطی با تابع هزینه‌ای مبتنی بر اطلاعات ماتریس ویژگی‌ها استفاده می‌شود تا رابطه‌ای صریح بین تمام نقاط اشکال برقرار شود. این استراتژی امکان انتخاب یکتا و منحصر به فرد نقاط بین مدل‌ها را

فراهم کرده و به ایجاد نگاشت نقطه به نقطه و تطبیق مورد نظر منجر می‌شود. برای ارزیابی شباهت دو مدل سه‌بعدی از تفاوت ویژگی‌های آن‌ها استفاده می‌شود. در اصل برای بیان اختلاف بین دو مدل، مجموع مجذورات اختلاف مؤلفه‌های متناظر ماتریس ویژگی‌ها، GD ، محاسبه می‌شود. بر این اساس، اختلاف بین مؤلفه i ام ماتریس ویژگی یک مدل با مؤلفه j ام مدل دیگر بر مبنای فاصله اقلیدسی و مطابق با رابطه (۵) محاسبه می‌شود.

$$Dist_{ij} = \sqrt{\sum(GD_i - GD_j)^2} ; i \in \{1,2, \dots, n_S\} \text{ and } j \in \{1,2, \dots, n_T\} \quad (5)$$

در رابطه (۵)، n_S و n_T به ترتیب تعداد رئوس مدل‌های سه‌بعدی منبع و هدف را نشان می‌دهند. مؤلفه‌های ماتریس $Dist_{ij}$ بر اساس فواصل ژئودزیکی هر نقطه تا نقاط برجسته (ویژه) محاسبه شده و تفاوت بین ویژگی‌های نقاط i و j در هر مدل را بیان می‌کنند. در ساده‌ترین حالت، کمینه مقدار این اختلاف، تطبیق بین نقاط i و j را نشان می‌دهد. با این حال، این روش تطبیق یکتا و منحصر به فرد بین نقاط مدل را تضمین نمی‌کند، زیرا ممکن است چندین نقطه در مدل هدف با یک نقطه در مدل منبع اختلاف فاصله یکسانی داشته باشند. برای ایجاد نگاشت نقطه‌ای بین مدل‌های ورودی، لازم است نگاشتی معکوس تعریف شود که نقاط را به‌طور یکتا تخصیص دهد. در روش پیشنهادی، این چالش با الهام از استراتژی حل مسئله تخصیص خطی برطرف شده است. در مسائل تخصیص خطی، هدف توزیع وظایف بین مؤلفه‌ها به گونه‌ای است که هزینه کل کمینه شود. برای محاسبه هزینه کل از ماتریس هزینه استفاده می‌شود که هزینه تخصیص هر مؤلفه از ردیف i به یک عمل در ستون j را مشخص می‌کند. مجموع هزینه‌های مؤلفه‌ها، هزینه کل را تعیین می‌کند. در این گونه مسائل، هر مؤلفه تنها یک وظیفه را انجام داده و هر وظیفه به یک مؤلفه اختصاص می‌یابد، به طوری که تمام وظایف با کمترین هزینه (مثلاً در کوتاه‌ترین زمان) انجام شوند. در روش پیشنهادی، ماتریس هزینه همان ماتریس $Dist_{ij}$ است. بنابراین، ساختار مسئله و مدل بهینه‌سازی می‌تواند در قالب معادله (۶) بیان شود.

$$\min Cost = \sum_{i=1}^n \sum_{j=1}^m Dist_{ij} \cdot a_{ij} \quad s.t. \quad \sum_{i=1}^n a_{ij} = 1 ; \sum_{j=1}^m a_{ij} = 1 ; a_{ij} \in \{0,1\} \quad (6)$$

معادله (۶) تابع هدف را بر حسب ماتریس $Dist_{ij}$ که از جنس فاصله است و قیده‌های باینری نگاشت (a_{ij}) تعریف می‌کند. این رویکرد تضمین می‌کند که تنها یک نقطه در شکل هدف برای نگاشت به نقطه‌ای در مدل منبع انتخاب شود. هدف این استراتژی، کمینه‌سازی هزینه از طریق انتخاب بهترین تطبیق است. تابع هدف و قیود آن، طبق معادله (۶)، نشان می‌دهند که نگاشت بهینه را می‌توان به صورت ریاضی با یک مسئله انتساب خطی بیان کرد که برای یافتن تطبیق بین نقاط به کار می‌رود. برای حل چنین مسائلی روش‌های متعددی وجود دارد، اما چالش اصلی، تعداد زیاد انتساب‌های مستقل ممکن در ماتریس هزینه است که این روش‌ها را پیچیده می‌کند. رایج‌ترین روش استفاده از الگوریتم کوهن-مانکراس^۱ (مجارستانی) است. الگوریتم‌های متنوعی از جمله نسخه‌های بهبود یافته مانکراس^۲ مبتنی بر گراف‌ها و ماتریس مجاورتی یا الگوریتم حراج^۳ با پیچیدگی‌های متفاوت برای حل مسئله تخصیص خطی ارائه شده‌اند. در این مقاله برای یافتن نقاط متناظر در مدل‌های سه‌بعدی منبع و هدف از الگوریتم کارآمد LAPJV استفاده شده است. این الگوریتم، یک پیاده‌سازی بهبود یافته از الگوریتم مانکراس است که با پیش‌پردازش ماتریس هزینه و بهره‌گیری از روش دایکسترا برای یافتن کوتاه‌ترین مسیرها، زمان محاسباتی را به‌طور قابل توجهی کاهش می‌دهد. الگوریتم LAPJV شامل دو فرآیند

¹ Kuhn-Munkres (KM)

² Improved-KM

³ Auction algorithm

اصلی است: فرآیند مقدماتی^۱ و فرآیند بهبود با اصلاحات اضافی^۲. در فرآیند مقدماتی، پیش‌پردازش ماتریس هزینه در سه مرحله انجام می‌شود: کاهش ستون^۳، انتقال کاهش^۴ و کاهش ردیف تقویتی^۵. مرحله کاهش ستون یا CRS شامل ساده‌سازی یا فشرده‌سازی اطلاعات ماتریس هزینه از طریق حذف ستون‌های غیر ضروری است که داده‌ها را برای انتساب مرتبط‌تر می‌کند و تحلیل ماتریس هزینه را تسهیل می‌نماید. مرحله انتقال کاهش یا RTS به انتقال مقادیر از یک بخش ماتریس هزینه به بخش دیگر اشاره دارد که معمولاً شامل تنظیم مقادیر برای بهینه‌سازی یا ساده‌سازی ساختار کلی و فرآیند محاسبه است. مرحله کاهش ردیف تقویتی یا ARRS با انجام عملیات سطری نظیر جمع یا تفریق مضربی از یک سطر به سطر دیگر، ماتریس تقویت شده به شکلی ساده‌تر تبدیل می‌کند تا حل مسئله تخصیص را امکان‌پذیر سازد. این فرآیند به شناسایی نقاطی کمک می‌کند که تمامی ارتباطات بین مؤلفه‌ها را به صورت همزمان برآورده می‌کنند. در نهایت، فرآیند بهبود با اصلاحات اضافی در الگوریتم LAPJV، نگاشتی دوطرفه ایجاد می‌کند (برای جزئیات بیشتر به [۶۸] مراجعه شود). الگوریتم^۳، مراحل یافتن نقاط متناظر بین دو مدل سه‌بعدی را بر اساس استراتژی حل مسئله تخصیص خطی با استفاده از LAPJV خلاصه می‌کند.

الگوریتم ۳: الگوریتم LAPJV

ورودی‌ها: - ماتریس هزینه ($Cost$) با ابعاد تعداد رئوس مدل‌ها ($Cost_{N \times N}$) حاوی مؤلفه‌های ماتریس‌های اختلاف ویژگی.

خروجی: - بردار ستونی انتساب‌ها ($assignment$).

۱: **فرآیند مقدماتی** ($Initialization Process$)

۲: **حلقه for** و تکرار آن به دفعات ابعاد ورودی

۳: اجرای گام کاهش ستون ($Column Reduction Step$)

۴: اجرای گام کاهش انتقال ($Reduction Transfer Step$)

۵: اجرای گام کاهش ردیف تقویتی ($Augmenting Row Reduction Step$)

۶: **فرآیند بهبود با اصلاحات اضافی** ($Augmentation Process$)

۷: **پایان حلقه for**

۸: تعیین مقادیر مطلوب هزینه‌ها ($Optimal Costs$)

۹: ذخیره انتساب‌های مطلوب (بهینه) در بردار $assignment$ و ارائه مقدار هزینه کمینه مربوط به انتساب‌های بهینه

۴- جزئیات پیاده‌سازی

روش پیشنهادی و الگوریتم‌های آن با استفاده از نرم افزار MATLAB پیاده سازی شده‌اند. تمامی آزمایش‌ها روی یک کامپیوتری ثابت با مشخصات سیستم عامل ۶۴ بیتی، پردازنده Core-i7 با سرعت ۲/۴ گیگاهرتز و ۸ گیگابایت حافظه انجام گرفته است. در فرآیند اجرای الگوریتم پیشنهادی از دو مجموعه داده استاندارد برای تناظریابی اشکال سه‌بعدی استفاده شده است. در ادامه، جزئیات و مشخصات هر یک از مجموعه داده‌ها و تنظیمات مورد نیاز برای اجرای صحیح الگوریتم‌ها ارائه شده است.

¹ Initialization process

² Augmentation process

³ Column Reduction Step

⁴ Reduction Transfer Step

⁵ Augmentation Row Reduction Step

۱-۴ - مجموعه داده‌ها

در این مقاله از مجموعه داده‌های پنج مارک DT4D^۱ و SMAL^۲ [۷۰] استفاده شده است. مجموعه داده DT4D شامل ۱۹۷۲ انیمیشن از اشکال مصنوعی انسان نما است که ۳۱ حیوان مختلف یا اشکال انسان‌نما با تراکم 4D را در بر می‌گیرد. ویژگی اصلی مجموعه داده DT4D، تغییرات حرکتی و غیرصُلب مدل‌های آن است. مجموعه داده SMAL شامل گونه‌های مختلف حیوانات چهارپا است که حاوی ۴۱ مدل از هشت گونه حیوانی در حالت‌ها و ژست‌های متنوع است. مجموعه داده SMAL شامل ۳۰۰ ژست منحصربه‌فرد از حیواناتی نظیر شیر، ببر، اسب، گاو، اسب آبی و سگ است. ویژگی عمده مجموعه داده SMAL، تغییرات شکل غیرایزومتریک بین گونه‌ها است که چالشی بزرگ برای روش‌های تناظریابی ایجاد می‌کند. از منظر آناتومیکی، امکان برقراری تناظر بین این اشکال وجود دارد، زیرا ساختارهای مشابهی بین مدل‌ها قابل تشخیص است. بنابراین، می‌توان تناظرهای معتبر به‌صورت خودکار برقرار کرد. در این مقاله از ۹ کلاس اشکال انسان‌نما از مجموعه داده DT4D و ۲۰ مدل از حیوانات چهارپا از مجموعه SMAL استفاده شده است.

۲-۴ - تنظیم پارامترها

در آزمایش‌های انجام شده برای روش پیشنهادی، تمامی پارامترها به‌صورت یکسان تنظیم شده‌اند. به‌طور خاص، از تنظیمات پیش‌فرض ارائه شده توسط نویسندگان روش‌های رقابتی برای متغیرهای مختلف استفاده شده است. اشکال ورودی با استفاده از الگوریتم LRVD [۷۱] دوباره مش‌بندی شده‌اند تا بایاس ناشی از مش‌های با اتصالات مشابه در مجموعه داده‌ها حذف شود. این مش‌بندی به‌صورت مستقل انجام شده است تا اشکال دارای اتصال یکسانی نباشند. به این ترتیب از اشکال با ۵۰۰۰ رأس استفاده شده است. در پیاده‌سازی روش‌های ارائه شده در [۵۰] و [۶۰] برای محاسبه نگاشت تابعی از ۱۲۰ تابع پایه ($k=120$) بهره گرفته شده است. این تعداد پایه برای مدل‌های غیرایزومتریک در مجموعه داده‌ها کافی بوده و اطلاعات لازم برای توصیف‌گرهای محاسباتی آن‌ها را فراهم می‌کند. برای تقریب دقیق تابع میانگین فاصله ژئودزیکی (AGD) و کاهش بار محاسباتی از ۱۶ نقطه نمونه ($p_s=16$) در شناسایی نقاط برجسته (ویژه) از اشکال ورودی استفاده شده است.

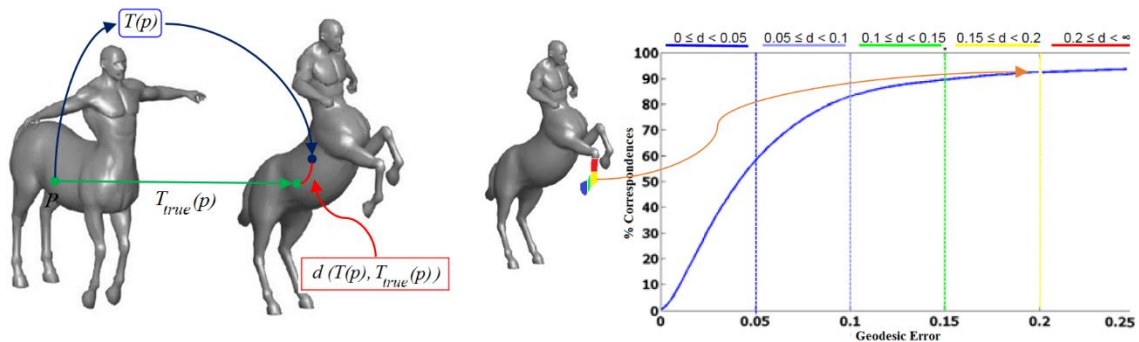
در ادامه، روش پیشنهادی با رویکردهای ارائه شده در [۴۸]، [۵۰] و [۶۰] مقایسه شده است که همگی از استراتژی مبتنی بر تطبیق ویژگی و توصیف‌گرها برای محاسبه تطبیق بین اشکال سه‌بعدی استفاده می‌کنند. به‌طور خلاصه، رویکرد [۴۸] از توصیف‌گرهای فاصله ژئودزیکی و پایه‌های تعریف شده برای تناظریابی در فضای متریک با حفظ اطلاعات شکلی بهره می‌برد. روش [۵۰]، محدودیت‌های حفظ توصیف‌گر و ویژگی‌های آن را در چارچوب نگاشت‌های تابعی ترکیب می‌کند تا تطبیق بهتری حاصل شود. روش [۶۰]، تناظرهای بین اشکال ورودی را با استفاده از توصیف‌گرهای تابعی و پایه‌های تعریف شده روی مدل‌های منبع و هدف محاسبه می‌کند. آزمایش‌های انجام شده و مقایسه نتایج، شامل رویکردهای مبتنی بر یادگیری ماشین نمی‌شوند، زیرا اغلب هزینه‌های آموزشی بالایی دارند و ممکن است به داده‌های دلخواه قابل تعمیم نباشند. در عوض، از تکنیک‌هایی استفاده شده است که نیازی به آموزش ندارند و به‌طور کلی برای مجموعه داده‌های مختلف قابل اجرا هستند. این رویکرد، امکان ارزیابی منصفانه کارایی روش پیشنهادی در مقایسه با عملکرد سایر روش‌های تطبیق شکل را فراهم می‌کند.

¹ Deforming Things4D

² Skinned Multi-Animal Linear model

۳-۴- معیارهای ارزیابی

برای ارزیابی نتایج تناظریابی روش پیشنهادی از پروتکل معرفی شده در [۷۲] استفاده شده است. این پروتکل کیفیت تطبیق را با استفاده از معیار خطای استاندارد (منحنی خطای ژئودزیکی) ارزیابی می‌کند. این منحنی، توزیع خطاهای تطبیق را نشان می‌دهد؛ به طوری که محور افقی (x) مقدار آستانه فاصله ژئودزیکی (خطای ژئودزیکی) و محور عمودی (y)، نرخ تطبیق یا میانگین درصد نقاطی را نشان می‌دهد که در فاصله مشخصی نسبت به مکان واقعی شان، gt ، پیش‌بینی شده‌اند. به عبارت دیگر برای هر نقطه از مدل منبع، فاصله ژئودزیکی بین نقطه پیش‌بینی شده و gt آن روی مدل هدف محاسبه می‌شود. این فاصله‌ها با ریشه دوم مساحت مدل هدف نرمال‌سازی می‌شوند. در شکل ۵ نمایش شهودی از پروتکل ارزیابی بین دو مدل فرضی ارائه شده است. نقاط سبز رنگ، نقاط متناظر در مدل منبع و هدف (به ترتیب در اشکال سه‌بعدی سمت چپ و راست) هستند؛ نقطه آبی نشان دهنده نقطه پیش‌بینی شده روی مدل هدف است و مسیر قرمز رنگ، فاصله ژئودزیکی بین نقاط آبی و سبز را روی مدل هدف نشان می‌دهد. هر چه این فاصله کمتر باشد، خطای پیش‌بینی کاهش می‌یابد. بر اساس منحنی ارائه شده در شکل ۶، میانگین نقاطی که در فاصله‌ای کمتر از 0.2 نسبت به gt آن‌ها روی مدل هدف پیش‌بینی شده‌اند، برابر با 90% درصد است. در این مثال، مقدار آستانه 0.2 در نظر گرفته شده است. همچنین برای ارزیابی بصری و مقایسه کیفی نتایج تطبیق از رنگ‌ها استفاده شده است. نقاط متناظر در اشکال منبع و هدف با رنگ‌های مشابه نشان داده می‌شوند.



شکل ۵. نمایش شهودی پروتکل ارزیابی و نمودار (منحنی) نرخ تطبیق.

۵- آزمایش‌ها و واکاوی نتایج

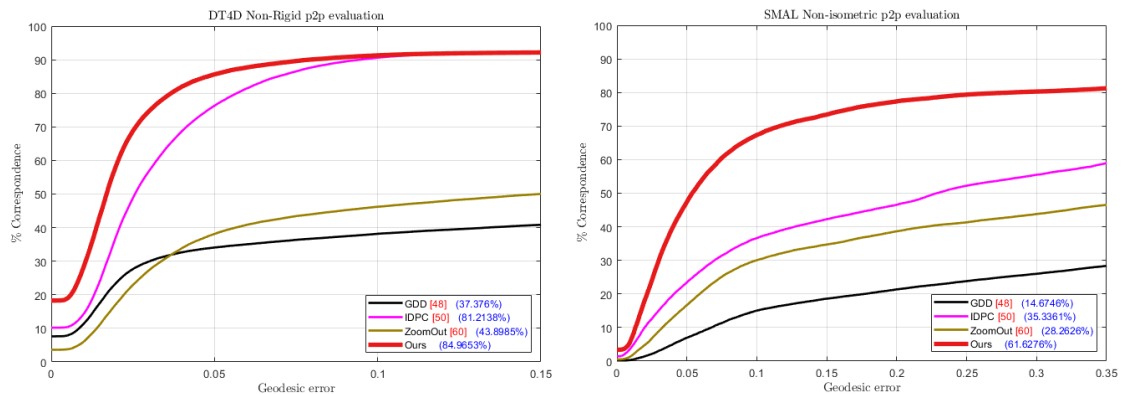
در این بخش، نتایج حاصل از روش پیشنهادی در تناظریابی اشکال سه‌بعدی بررسی شده و نقاط قوت و ضعف آن تحلیل می‌شود. نتایج کمی و کیفی روش پیشنهادی با رویکردهای $[48]$ ، $[50]$ و $[60]$ روی مجموعه داده‌های DT4D و SMAL مقایسه شده است. شکل ۶ منحنی‌های مقایسه‌ای را بر اساس پروتکل ارزیابی [۷۲] برای روش‌های مختلف روی این مجموعه داده‌ها نمایش می‌دهد. در این منحنی‌ها، درصد تناظرهای پیش‌بینی شده در فاصله ژئودزیکی مشخص نسبت به حقیقت پایه (gt) ترسیم شده است. در شکل ۶ (سمت چپ)، منحنی‌های روش پیشنهادی با رویکردهای $[48]$ ، $[50]$ و $[60]$ روی مجموعه داده غیرصلب DT4D مقایسه شده‌اند. نمودارها نشان می‌دهند که روش پیشنهادی عملکرد بهتری نسبت به سایر روش‌ها دارد. برای مثال، درصد تناظرهای پیش‌بینی شده با روش پیشنهادی در آستانه فاصله ژئودزیکی 0.05 ، بالاتر از سایر منحنی‌ها و به طور تقریبی برابر با 85% است. همچنین، بیشترین سطح زیر منحنی معادل $84/9653\%$ با روش پیشنهادی به دست آمده است. به طور میانگین، تعداد نقاط پیش‌بینی شده صحیح با روش

¹ Ground truth

پیشنهادی روی مدل‌های مجموعه داده DT4D از سایر روش‌ها بیشتر است. به‌طور مشابه در شکل ۶ (سمت راست)، منحنی‌های مقایسه‌ای روی مجموعه داده‌های غیرایزومتریک SMAL ترسیم شده‌اند. نتایج نشان می‌دهند که بیشترین سطح زیر منحنی برابر با ۶۱/۶۲۷۶٪ متعلق به روش پیشنهادی است. به عبارت دیگر، میانگین تعداد نقاط پیش‌بینی شده صحیح با روش پیشنهادی روی مدل‌های مجموعه داده SMAL نسبت به سایر روش‌ها بیشتر است. در مجموع، نتایج آزمایش‌ها نشان می‌دهند که روش پیشنهادی در مقایسه با رویکردهای [۴۸]، [۵۰] و [۶۰] در هر دو مجموعه داده عملکرد بهتری داشته و نرخ تطبیق بالاتری را کسب کرده است. در ادامه، نتایج کمی برای مدل‌های سه‌بعدی مختلف از مجموعه داده‌ها در جدول ۱ ارائه شده است. معیارهای ارزیابی شامل میانگین خطای ژئودزیک و میانگین زمان اجرا (برحسب ثانیه) برای روش‌های مختلف است که به ترتیب با نمادهای Av. Geo. Err. و Av. Time نشان داده شده‌اند. نتایج نشان می‌دهند که روش پیشنهادی در مقایسه با رویکردهای [۴۸]، [۵۰] و [۶۰] به مقادیر خطای کمتری دست یافته است. به‌طور خاص، میانگین خطای ژئودزیک روش پیشنهادی برای مجموعه داده‌های DT4D و SMAL به ترتیب (۰/۰۷۱) و (۰/۱۷۱۸) محاسبه شده است که کمترین مقدار خطا در مقایسه با سایر روش‌ها است. همچنین از نظر زمان اجرا، روش پیشنهادی با روش [۵۰] قابل مقایسه بوده و نسبت به سایر روش‌ها سریع‌تر عمل می‌کند. بر اساس نتایج، روش پیشنهادی در مجموعه داده SMAL که شامل مدل‌های غیرایزومتریک است، از نظر کمی و کیفی عملکرد بهتری نسبت به سایر روش‌ها دارد. شکل ۷ ارزیابی بصری و مقایسه کیفی نتایج تناظریابی روش‌های مختلف بین جفت مدل‌های ورودی را نمایش می‌دهد. برتری کیفی نتایج روش پیشنهادی نسبت به رویکردهای [۴۸]، [۵۰] و [۶۰] با استفاده از نمایش رنگی برای هر مجموعه داده نشان داده شده است.

جدول ۱: نتایج ارزیابی کمی در مجموعه داده‌های مختلف و مقایسه روش پیشنهادی با رویکردهای [۴۸]، [۵۰] و [۶۰].

روش	مجموعه داده	Av. Geo. Err.	Av. Time in sec
GDD [۴۸]	DT4D	۰/۳۶۵۸	۴۹/۰۰
	SMAL	۰/۵۷۲۵	۵۶/۵۰
IDPC [۵۰]	DT4D	۰/۰۸۲۰	۱۷/۱۳
	SMAL	۰/۴۵۰۶	۱۶/۵۱
Zoom Out [۶۰]	DT4D	۰/۳۳۵۴	۲۷/۹۱
	SMAL	۰/۴۸۷۳	۲۷/۱۰
روش پیشنهادی	DT4D	۰/۰۷۱۰	۲۱/۳۱
	SMAL	۰/۱۷۱۸	۲۲/۱۹

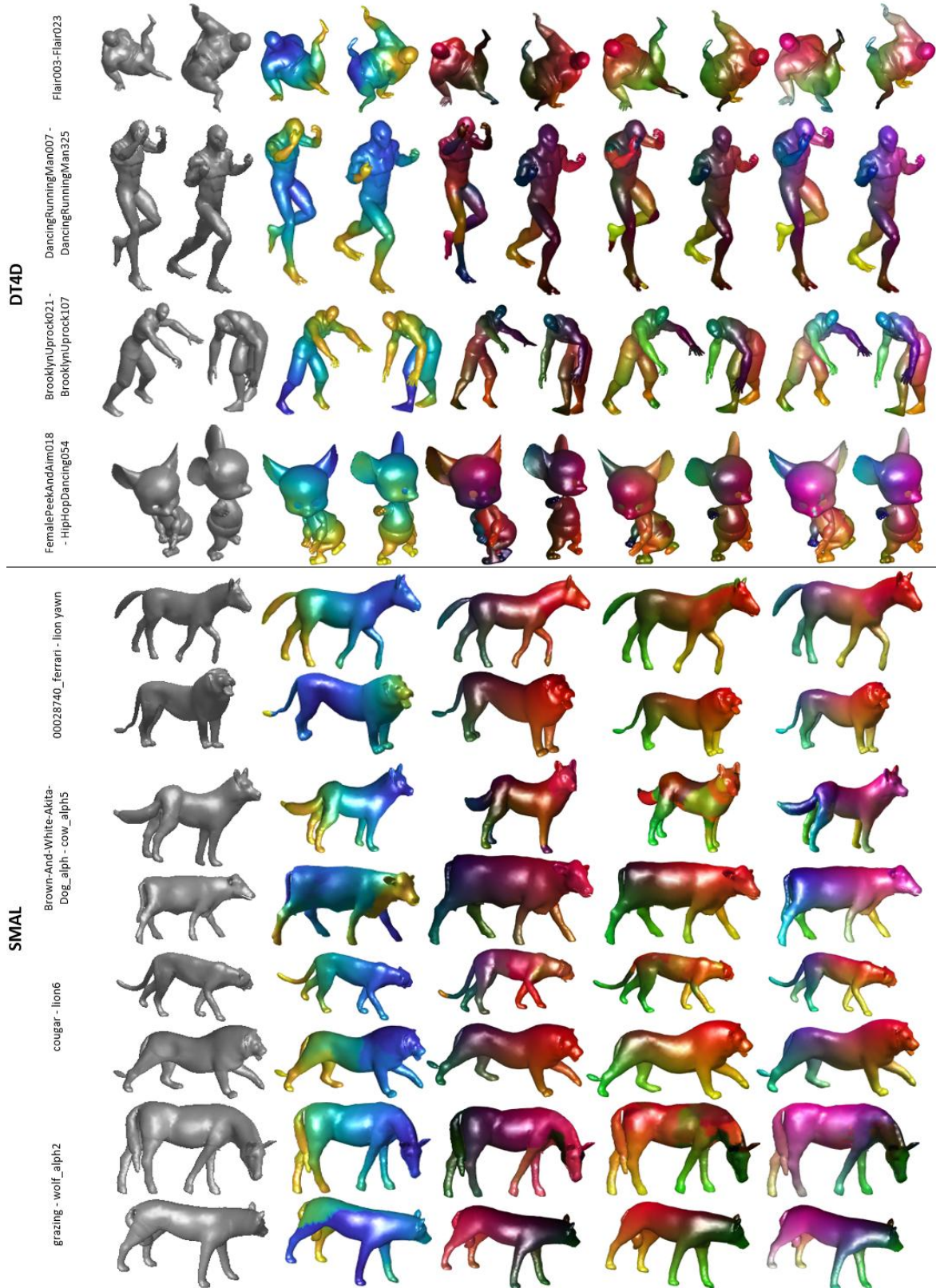


شکل ۶: نتایج کمی و مقایسه روش پیشنهادی با رویکردهای [۴۸]، [۵۰] و [۶۰] بر اساس پروتکل ارزیابی [۷۲].

۱-۵- پیچیدگی محاسباتی

در این بخش، پیچیدگی محاسباتی روش پیشنهادی برای یافتن تناظر بین دو مدل سه‌بعدی تحلیل شده است. الگوریتم پیشنهادی در سه مرحله اصلی اجرا می‌شود و پیچیدگی محاسباتی هر مرحله به صورت جداگانه تخمین زده شده است. در مرحله اول برای محاسبه فواصل ژئودزیکی روی مدل‌ها از الگوریتم Fast Marching با پیچیدگی محاسباتی $O(n \log n)$ استفاده شده است که n تعداد نقاط (رئوس) مدل‌ها است. محاسبات فاصله‌های ژئودزیکی، p_s بار برای شناسایی اولین نقطه برجسته (ویژه) و $(n_{sp} - 1)$ بار برای سایر نقاط برجسته انجام می‌شود. بنابراین، پیچیدگی محاسباتی مرحله اول برای شناسایی نقاط برجسته مدل برابر است با: $O(p_s \cdot n \log n + (n_{sp} - 1) n \log n)$. فشرده‌ترین بخش مرحله دوم از نظر محاسباتی به خط شماره چهارم الگوریتم ۲ مربوط می‌شود که به روش محاسبه فواصل ژئودزیکی روی سطح مش وابسته است. در این مرحله، استخراج ماتریس ویژگی مدل‌ها بر اساس فواصل ژئودزیکی نقاط از نقاط ویژه با استفاده از الگوریتم Fast Marching انجام می‌شود که پیچیدگی محاسباتی آن $O(n \log n)$ است. مرحله سوم که به برقراری تطبیق بین نقاط مدل‌ها با الگوریتم ۳ (LAPJV) اختصاص دارد، به دلیل عملیات سطری و ستونی روی ماتریس ورودی با ابعاد $n \times n$ ، دارای پیچیدگی محاسباتی $O(n^2)$ است. برای مطالعه دقیق‌تر تحلیل پیچیدگی الگوریتم ۳، خواننده می‌تواند به بخش ضمیمه (الف) مراجعه کند. به‌طور خلاصه، پیچیدگی محاسباتی مراحل مختلف روش پیشنهادی در جدول ۲ ارائه شده است. در حالت کلی، پیچیدگی کل روش پیشنهادی از مجموع پیچیدگی مراحل آن محاسبه می‌شود که در بدترین حالت برابر با $O(n^2)$ است.

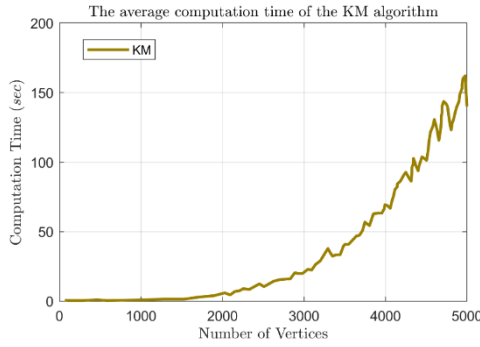
در این مقاله، برای تسریع الگوریتم LAPJV از یک فایل mex نوشته شده در زبان برنامه‌نویسی C++ استفاده شده است. به‌منظور نشان دادن کارایی الگوریتم LAPJV در کاهش قابل توجه زمان اجرا از الگوریتم‌های مشابه مانند الگوریتم حراج [۷۳]، کوهن-مانکراس (KM) و نسخه بهبود یافته آن [۷۴] به جای LAPJV استفاده شد. برای مقایسه زمان محاسباتی الگوریتم‌ها، نتایج روش پیشنهادی روی ۱۰۰ جفت مدل با تعداد رئوس بین ۱۰ تا ۵۰۰۰ بررسی شده است. در شکل ۸، نمودارهای مقایسه‌ای زمان‌های محاسباتی الگوریتم‌ها با تغییر تعداد رئوس مدل ارائه شده است که به ترتیب میانگین زمان‌های محاسباتی برای الگوریتم LAPJV، KM، بهبود یافته KM و الگوریتم حراج را نشان می‌دهد. به‌عنوان مثال، شکل‌های ۸-الف تا ۸-ج نشان می‌دهند که برای مدل‌هایی با ۳۷۰۰ نقطه، میانگین زمان محاسباتی الگوریتم LAPJV کمتر از ۲۰ ثانیه است. در حالی که، میانگین زمان الگوریتم حراج با ۸۰۰ نقطه، الگوریتم KM با ۲۷۰۰ نقطه و بهبود یافته KM با ۲۹۰۰ نقطه از ۲۰ ثانیه فراتر می‌رود. همچنین در شکل ۸-د مشاهده می‌شود که الگوریتم LAPJV با تغییر تعداد نقاط مدل از ۱۰ تا ۵۰۰۰، همواره زمان محاسباتی بسیار کمتری نسبت به سه الگوریتم دیگر دارد. اگرچه میانگین با تغییر تعداد نقاط مدل (رأس‌های شکل سه‌بعدی)، میانگین زمان محاسباتی الگوریتم LAPJV افزایش می‌یابد، اما همچنان در حدود ۲۰ ثانیه باقی می‌ماند.



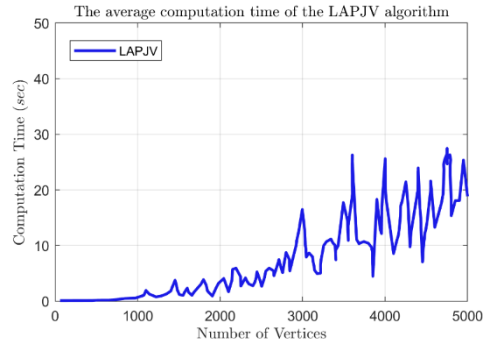
شکل ۷. نمایش کیفی نتایج روش پیشنهادی با رویکردهای [۴۸]، [۵۰] و [۶۰] روی اشکال مجموعه داده‌های SMAL و DT4D

جدول ۲: پیچیدگی محاسباتی روش پیشنهادی (n = تعداد رئوس، n_{sp} = تعداد نقاط برجسته)

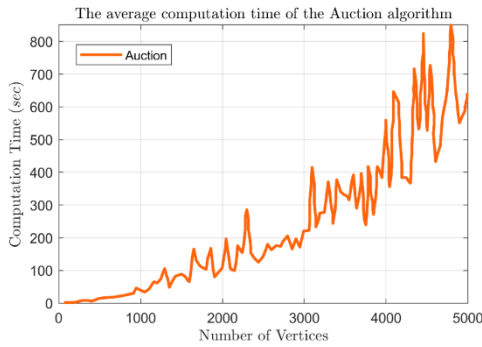
مراحل	هدف الگوریتم در هر مرحله	پیچیدگی محاسباتی
۱	شناسایی نقاط برجسته (ویژه)	$O(p_s n \log n + (n_{sp} - 1) n \log n)$
۲	استخراج ماتریس ویژگی	$O(n \log n)$
۳	برقراری تطابق بین نقاط با الگوریتم LAPJV	$O(n^2)$



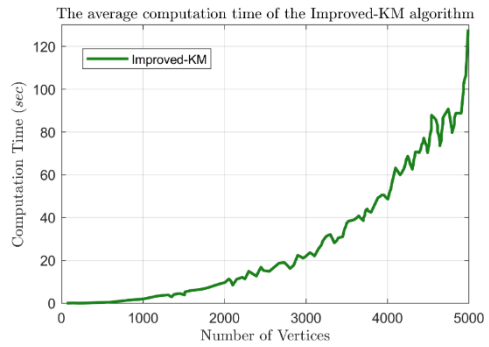
(ب)



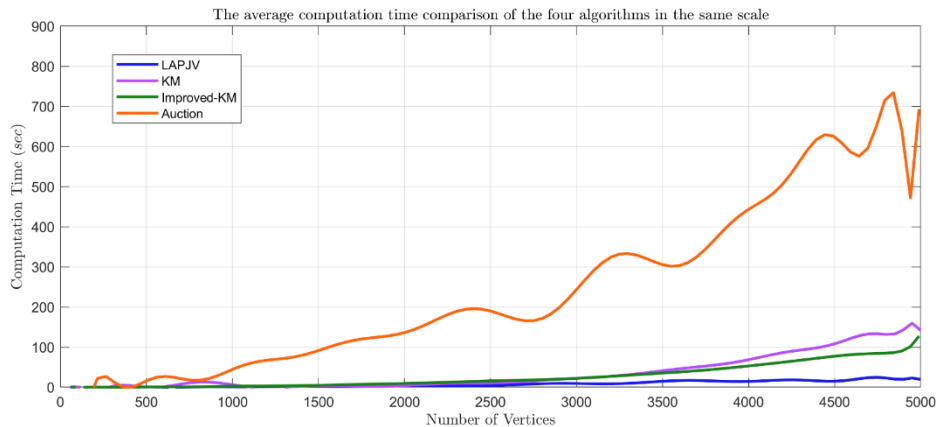
(الف)



(ج)



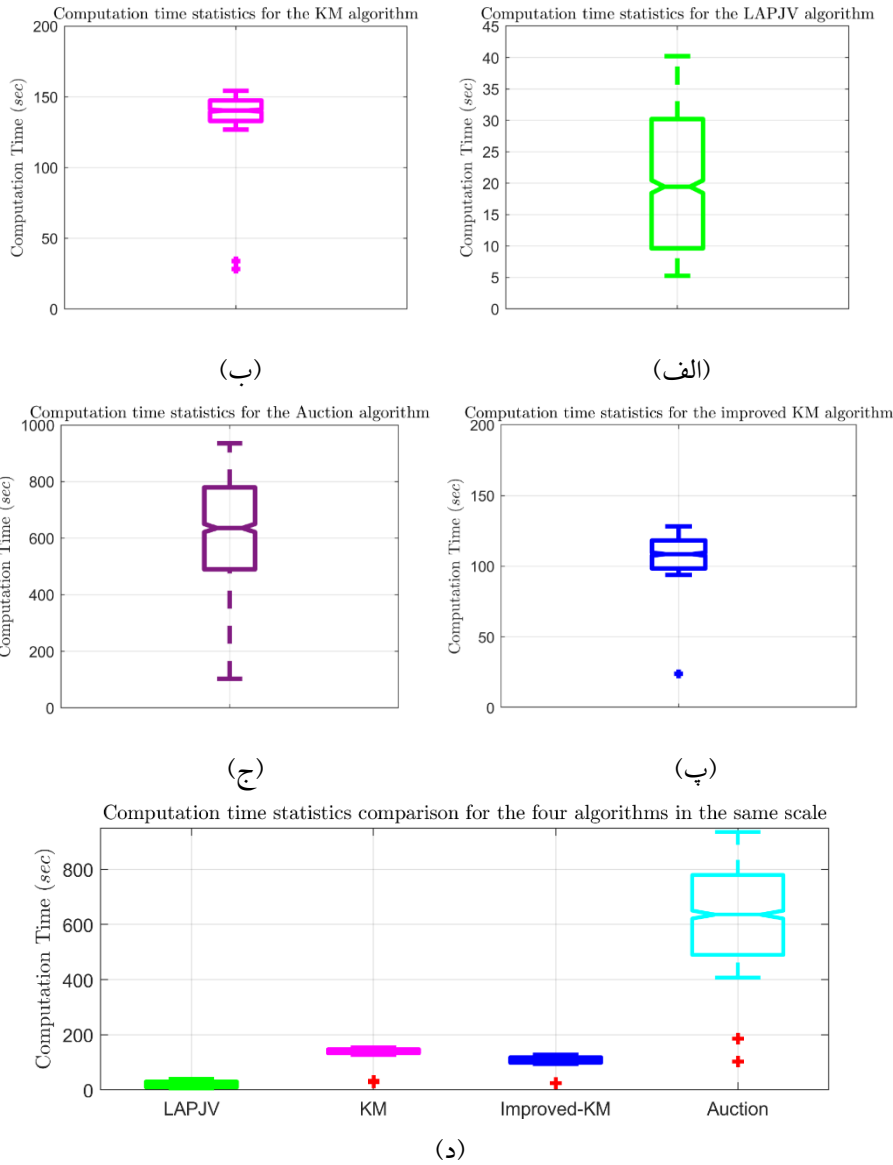
(پ)



(د)

شکل ۸. نمودارهای مقایسه‌ای زمان محاسباتی بین الگوریتم‌های KM، LAPJV، KM، بهبود یافته KM و الگوریتم حراج با تغییر نقاط مدل از ۱۰ تا ۵۰۰۰. (الف) میانگین زمان محاسباتی الگوریتم LAPJV. (ب) میانگین زمان محاسباتی الگوریتم KM. (پ) میانگین زمان محاسباتی الگوریتم بهبود یافته KM. (ج) میانگین زمان محاسباتی الگوریتم حراج. (د) مقایسه میانگین زمان محاسباتی چهار الگوریتم در یک مقیاس.

برای تحلیل دقیق‌تر، آمار زمان محاسبه برای حالت ۵۰۰۰ نقطه از ۱۰۰ جفت مدل (شامل تمام رئوس مدل‌ها) پیاده‌سازی شده و نتایج آزمایش در شکل ۹ ارائه شده است. میانگین زمان محاسباتی الگوریتم LAPJV برابر با ۲۰ ثانیه است. در حالی که این مقدار برای الگوریتم KM برابر با ۱۴۰ ثانیه، برای الگوریتم بهبود یافته KM برابر با ۱۰۸ ثانیه و الگوریتم حراج بیش از ۶۰۰ ثانیه است. به عبارت دیگر، الگوریتم LAPJV زمان محاسباتی را به ترتیب ۸۵/۷۱٪ نسبت به الگوریتم KM، ۸۱/۴۸٪ نسبت به الگوریتم بهبود یافته KM و بیش از ۹۶٪ نسبت به الگوریتم حراج کاهش می‌دهد.



شکل ۹. مقایسه آماری زمان محاسباتی بین الگوریتم LAPJV، KM، بهبود یافته KM و الگوریتم حراج برای مدل با ۵۰۰۰ رأس. (الف) آمار زمان محاسباتی الگوریتم LAPJV. (ب) آمار زمان محاسباتی الگوریتم KM. (پ) آمار زمان محاسباتی الگوریتم بهبود یافته KM. (ج) آمار زمان محاسباتی الگوریتم حراج. (د) مقایسه آماری زمان محاسباتی چهار الگوریتم در یک مقیاس.

در جدول ۳، نتایج آزمایش‌ها به صورت خلاصه ارائه شده است. نتایج نشان می‌دهند که زمان محاسباتی الگوریتم LAPJV در مقایسه با الگوریتم های KM، حراج و بهبود یافته KM به طور قابل توجهی کمتر بوده و عملکرد سریع تری دارد.

جدول ۳: پیچیدگی محاسباتی و زمان اجرای مرحله ۳ از روش پیشنهادی (n = تعداد رئوس)

نام الگوریتم برای حل مسایل انتساب خطی	پیچیدگی محاسباتی	متوسط زمان اجرا با $n = 5k$ (بر حسب ثانیه)
الگوریتم حراج (Auction) [۷۳]	$O(n^2 \log n)$	$>> 140$
الگوریتم کلاسیک کوهن-مانکراس [۷۴]	$O(n^4)$	۱۴۰
الگوریتم بهبود یافته کوهن-مانکراس [۷۴]	$O(n^3)$	۱۰۸
الگوریتم LAPJV	$O(n^2)$	۲۰

۲-۵ - مقایسه مدت زمان اجرا

زمان محاسباتی هر الگوریتم به طور مستقیم با پیچیدگی محاسباتی آن مرتبط است. بنابراین، زمان محاسباتی روش پیشنهادی با روش‌های رقابتی [۴۸]، [۵۰] و [۶۰] مقایسه شده است. با توجه به سخت افزار معرفی شده در این تحقیق و شبیه‌سازی‌های انجام شده، زمان محاسباتی روش‌ها اندازه‌گیری شده است. در جدول ۱، زمان اجرای روش‌های تطبیق ارائه شده است. نتایج نشان می‌دهند که روش پیشنهادی از نظر زمان محاسباتی با اختلاف ۵ ثانیه نسبت به روش [۵۰] در رتبه دوم قرار دارد. به طور میانگین، زمان مورد نیاز برای یافتن تطبیق بین دو مدل سه‌بعدی از مجموعه داده‌های مختلف با روش [۴۸] برابر با ۵۳ ثانیه، با روش [۶۰] برابر با ۲۷ ثانیه و با روش پیشنهادی برابر با ۲۲ ثانیه است. همچنین، مقایسه زمان محاسباتی روش پیشنهادی با رویکردهای [۴۸]، [۵۰] و [۶۰] براساس مفهوم $Big-O$ نشان می‌دهد که در بدترین حالت، پیچیدگی محاسباتی روش پیشنهادی برابر با $O(n^2)$ است که کمتر از پیچیدگی محاسباتی $O(n^3)$ برای روش‌های [۴۸]، [۵۰] و [۶۰] است که n تعداد رئوس مدل‌ها است.

۳-۵ - محدودیت‌ها و کارهای آینده

در این تحقیق فرض شده است که اشکال ورودی به طور نسبی منظم بوده و دارای خصایص مشابهی هستند. به طور خاص، انتظار می‌رود که دست کم ساختارهای مشابه در مقیاس بزرگ بین اشکالی که باید تطبیق داده شوند، قابل مشاهده باشد. علاوه بر این، روش پیشنهادی از مدل‌های سه‌بعدی کامل استفاده می‌کند و اشکال جزئی یا مدل‌های تخریب شده در نظر گرفته نمی‌شوند. روش پیشنهادی با استفاده از حل مسئله انتساب خطی LAPJV به نتایج مطلوبی روی مجموعه داده‌ها دست یافته است. همان‌طور که منحنی‌های خطای ژئودزیک در شکل ۶ برای مجموعه داده SMAL نشان می‌دهند روش پیشنهادی می‌تواند با انتخاب ویژگی‌های هندسی جدید و ترکیب آن‌ها با ویژگی فعلی از طریق تکنیک‌های سریع در جستجوی نقاط متناظر یا افزودن مرحله پس پردازش بهبود یابد. مورد اول از پیش‌بینی نقاط پرت به عنوان نقاط متناظر جلوگیری کرده یا ابهامات ناشی از تقارن را رفع می‌کند. مورد دوم به افزایش سرعت اجرای الگوریتم LAPJV کمک می‌کند. مورد سوم برای بهبود نگاشت پیش‌بینی شده به کار می‌رود.

۶- نتیجه‌گیری

در این مقاله، روشی کارآمد مبتنی بر استراتژی انتساب‌های خطی با استفاده از الگوریتم LAPJV برای یافتن تطبیق منحصر به فرد بین نقاط یک جفت شکل سه‌بعدی پیشنهاد شده است. تابع هزینه مسئله تخصیص خطی بر اساس شناسایی نقاط برجسته اشکال ورودی و استخراج ماتریس ویژگی‌ها با استفاده از توصیفگر فاصله ژئودزیک تعریف شده است. مزایای روش پیشنهادی در مقایسه با روش‌های دیگر روی مجموعه داده‌های بنچ مارک شناخته شده، ارزیابی شده است. نتایج نشان می‌دهند که روش پیشنهادی برای اشکال غیر صُلب مجموعه داده DT4D و اشکال غیر ایزومتریک مجموعه داده SMAL در مقایسه با رویکردهای [۴۸]، [۵۰] و [۶۰] نرخ تطبیق بالاتری دارد. همان‌طور که در نمودارهای شکل ۶ نشان داده شده است، میانگین نرخ تطبیق برای آستانه خطای ژئودزیک کمتر از ۰/۰۵ و ۰/۱۵ به ترتیب ۰/۸۵ و ۰/۷۳ است که متعلق به روش پیشنهادی است. همچنین بر اساس مقادیر جدول ۱، کمترین میانگین خطای ژئودزیک برای مجموعه داده‌های SMAL و DT4D به روش پیشنهادی تعلق دارد. علاوه بر این، مقایسه‌های انجام شده از نظر زمان اجرا و پیچیدگی محاسباتی، برتری روش پیشنهادی را نسبت به رویکردهای مقایسه شده نشان می‌دهد. به‌طور کلی، روش پیشنهادی در مقایسه با رویکردهای [۴۸]، [۵۰] و [۶۰] به بهبود نرخ تطبیق، کاهش پیچیدگی محاسباتی و افزایش سرعت اجرا دست یافته است. در آینده، قصد داریم چارچوب روش پیشنهادی را با در نظر گرفتن ملاحظات بیشتر به طیف گسترده‌تری از اشکال ورودی گسترش دهیم تا امکان استفاده از آن برای برقراری تطبیق دقیق‌تر بین مدل‌های سه‌بعدی در سایر مجموعه داده‌ها با ویژگی‌های متنوع، اشکال جزئی یا تخریب شده یا مدل‌هایی با تغییرات شکلی زیاد فراهم شود و نقاط ضعف روش پیشنهادی در این موارد برطرف شود.

مراجع

- [1] O. van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or, "A Survey on Shape Correspondence," *Computer Graphics Forum*, vol. 30, no. 6, pp. 1681-1707, 2011, doi: <https://doi.org/10.1111/j.1467-8659.2011.01884.x>.
- [2] R. M. Dyke *et al.*, "SHREC'20: Shape correspondence with non-isometric deformations," *Computers & Graphics*, vol. 92, pp. 28-43, 2020/11/01/ 2020, doi: <https://doi.org/10.1016/j.cag.2020.08.008>.
- [3] A. Sedaghat and H. Ebadi, "Remote Sensing Image Matching Based on Adaptive Binning SIFT Descriptor," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 10, pp. 5283-5293, 2015, doi: 10.1109/TGRS.2015.2420659.
- [4] D. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Pearson, 2012.
- [5] M. Mirloo and H. Ebrahimnezhad, "Non-rigid 3D object retrieval using directional graph representation of wave kernel signature," *Multimedia Tools and Applications*, vol. 77, no. 6, pp. 6987-7011, 2018/03/01 2018, doi: 10.1007/s11042-017-4617-x.
- [6] R. M. Dyke, Y.-K. Lai, P. L. Rosin, and G. K. L. Tam, "Non-Rigid Registration under Anisotropic Deformations," *Comput. Aided Geom. Des.*, vol. 71, no. C, pp. 142-156, 2019, doi: 10.1016/j.cagd.2019.04.014.
- [7] T. Jiang *et al.*, "Huber- \mathbb{S}^L_1 -based non-isometric surface registration," *The Visual Computer*, vol. 35, no. 6, pp. 935-948, 2019/06/01 2019, doi: 10.1007/s00371-019-01670-1.
- [8] D. Ezuz, B. Heeren, O. Azencot, M. Rumpf, and M. Ben-Chen, "Elastic Correspondence between Triangle Meshes," *Computer Graphics Forum*, vol. 38, no. 2, pp. 121-134, 2019, doi: <https://doi.org/10.1111/cgf.13624>.
- [9] G. K. Tam *et al.*, "Registration of 3D point clouds and meshes: A survey from rigid to nonrigid," *IEEE transactions on visualization and computer graphics*, vol. 19, no. 7, pp. 1199-1217, 2012.
- [10] Y. Sahillioğlu, "Recent advances in shape correspondence," *The Visual Computer*, vol. 36, no. 8, pp. 1705-1721, 2020.
- [11] J. Pokrass, A. M. Bronstein, M. M. Bronstein, P. Sprechmann, and G. Sapiro, "Sparse Modeling of

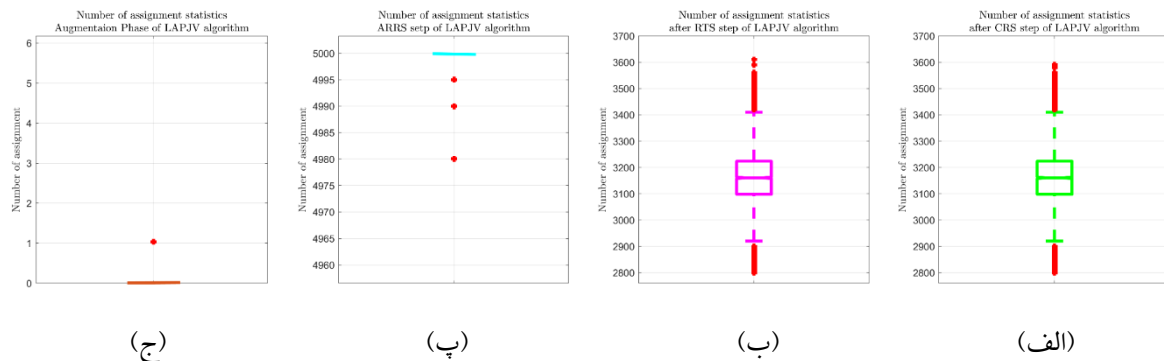
- Intrinsic Correspondences," *Computer Graphics Forum*, vol. 32, no. 2pt4, pp. 459-468, 2013, doi: <https://doi.org/10.1111/cgf.12066>.
- [12] O. van Kaick, H. Zhang, and G. Hamarneh, "Bilateral Maps for Partial Matching," *Computer Graphics Forum*, vol. 32, no. 6, pp. 189-200, 2013, doi: <https://doi.org/10.1111/cgf.12084>.
- [13] M. Ovsjanikov, M. Ben-Chen, J. Solomon, A. Butscher, and L. Guibas, "Functional Maps: A Flexible Representation of Maps between Shapes," *ACM Trans. Graph.*, vol. 31, no. 4, 2012, doi: 10.1145/2185520.2185526.
- [14] Y. Kleiman and M. Ovsjanikov, "Robust Structure-Based Shape Correspondence," *Computer Graphics Forum*, vol. 38, no. 1, pp. 7-20, 2019, doi: <https://doi.org/10.1111/cgf.13389>.
- [15] D. Nogneng and M. Ovsjanikov, "Informative Descriptor Preservation via Commutativity for Shape Matching," *Comput. Graph. Forum*, vol. 36, no. 2, pp. 259-267, 2017, doi: 10.1111/cgf.13124.
- [16] R. Marin, S. Melzi, E. Rodola, and U. Castellani, "Farm: Functional automatic registration method for 3d human bodies," in *Computer Graphics Forum*, 2020, vol. 39, no. 1: Wiley Online Library, pp. 160-173.
- [17] A. Gehre, M. Bronstein, L. Kobbelt, and J. Solomon, "Interactive curve constrained functional maps," in *Computer Graphics Forum*, 2018, vol. 37, no. 5: Wiley Online Library, pp. 1-12.
- [18] J. Sun, M. Ovsjanikov, and L. Guibas, "A concise and provably informative multi-scale signature based on heat diffusion," in *Computer graphics forum*, 2009, vol. 28, no. 5: Wiley Online Library, pp. 1383-1392.
- [19] M. Aubry, U. Schlickewei, and D. Cremers, "The wave kernel signature: A quantum mechanical approach to shape analysis," in *2011 IEEE international conference on computer vision workshops (ICCV workshops)*, 2011: IEEE, pp. 1626-1633.
- [20] H. Chui and A. Rangarajan, "A new algorithm for non-rigid point matching," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, 2000, vol. 2: IEEE, pp. 44-51.
- [21] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 11, no. 6, pp. 567-585, 1989.
- [22] R. W. Sumner and J. Popović, "Deformation transfer for triangle meshes," *ACM Transactions on graphics (TOG)*, vol. 23, no. 3, pp. 399-405, 2004.
- [23] N. Aigerman, R. Poranne, and Y. Lipman, "Lifted bijections for low distortion surface mappings," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, pp. 1-12, 2014.
- [24] O. Weber and D. Zorin, "Locally injective parametrization with arbitrary fixed boundaries," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, pp. 1-12, 2014.
- [25] N. Aigerman and Y. Lipman, "Orbifold tutte embeddings," *ACM Trans. Graph.*, vol. 34, no. 6, pp. 190:1-190:12, 2015.
- [26] N. Aigerman and Y. Lipman, "Hyperbolic orbifold tutte embeddings," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 217:1-217:14, 2016.
- [27] N. Aigerman, R. Poranne, and Y. Lipman, "Seamless surface mappings," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, pp. 1-13, 2015.
- [28] D. Panozzo, I. Baran, O. Diamanti, and O. Sorkine-Hornung, "Weighted averages on surfaces," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, pp. 1-12, 2013.
- [29] D. Ezuz, J. Solomon, and M. Ben-Chen, "Reversible harmonic maps between discrete surfaces," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 2, pp. 1-12, 2019.
- [30] P. Schmidt, M. Campen, J. Born, and L. Kobbelt, "Inter-surface maps via constant-curvature metrics," *ACM Transactions on Graphics (TOG)*, vol. 39, no. 4, pp. 119: 1-119: 15, 2020.
- [31] M. Ovsjanikov et al., "Computing and processing correspondences with functional maps," in *SIGGRAPH ASIA 2016 Courses*, 2016, pp. 1-60.
- [32] Y. Aflalo, H. Brezis, and R. Kimmel, "On the optimality of shape and data representation in the spectral domain," *SIAM Journal on Imaging Sciences*, vol. 8, no. 2, pp. 1141-1160, 2015.
- [33] A. Kovnatsky, K. Glashoff, and M. M. Bronstein, "MADMM: a generic algorithm for non-smooth optimization on manifolds," in *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V 14*, 2016: Springer, pp. 680-696.
- [34] Y. Choukroun, A. Shtern, A. Bronstein, and R. Kimmel, "Hamiltonian operator for spectral shape

- analysis," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 2, pp. 1320-1331, 2018.
- [35] S. Melzi, E. Rodolà, U. Castellani, and M. M. Bronstein, "Localized manifold harmonics for spectral shape analysis," in *Computer Graphics Forum*, 2018, vol. 37, no. 6: Wiley Online Library, pp. 20-34.
- [36] S. Melzi, R. Marin, P. Musoni, F. Bardoni, M. Tarini, and U. Castellani, "Intrinsic/extrinsic embedding for functional remeshing of 3D shapes," *Computers & Graphics*, vol. 88, pp. 1-12, 2020.
- [37] G. Patané, "Laplacian spectral basis functions," *Computer aided geometric design*, vol. 65, pp. 31-47, 2018.
- [38] L. Cosmo, M. Panine, A. Rampini, M. Ovsjanikov, M. M. Bronstein, and E. Rodolà, "Isospectralization, or how to hear shape, style, and correspondence," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7529-7538.
- [39] M. Eisenberger, Z. Löhner, and D. Cremers, "Divergence-Free Shape Correspondence by Deformation," in *Computer Graphics Forum*, 2019, vol. 38, no. 5: Wiley Online Library, pp. 1-12.
- [40] E. Küpçü and Y. Yemez, "Diffusion-based isometric depth correspondence," *Computer Vision and Image Understanding*, vol. 189, p. 102808, 2019.
- [41] M. M. Bronstein and I. Kokkinos, "Scale-invariant heat kernel signatures for non-rigid shape recognition," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 13-18 June 2010 2010, pp. 1704-1711, doi: 10.1109/CVPR.2010.5539838.
- [42] F. Tombari, S. Salti, and L. Di Stefano, "Unique Signatures of Histograms for Local Surface Description," in *Computer Vision – ECCV 2010*, Berlin, Heidelberg, K. Daniilidis, P. Maragos, and N. Paragios, Eds., 2010// 2010: Springer Berlin Heidelberg, pp. 356-369.
- [43] D. Boscaini, J. Masci, E. Rodolà, M. M. Bronstein, and D. Cremers, "Anisotropic Diffusion Descriptors," *Computer Graphics Forum*, vol. 35, no. 2, pp. 431-441, 2016, doi: <https://doi.org/10.1111/cgf.12844>.
- [44] L. Cosmo, E. Rodola, M. M. Bronstein, A. Torsello, D. Cremers, and Y. Sahillioglu, "SHREC'16: Partial matching of deformable shapes," *Proc. 3DOR*, vol. 2, no. 9, p. 12, 2016.
- [45] M. Vestner *et al.*, "Efficient Deformable Shape Correspondence via Kernel Matching," in *2017 International Conference on 3D Vision (3DV)*, 10-12 Oct. 2017 2017, pp. 517-526, doi: 10.1109/3DV.2017.00065.
- [46] R. R. Coifman *et al.*, "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps," *Proceedings of the National Academy of Sciences*, vol. 102, no. 21, pp. 7426-7431, 2005, doi: doi:10.1073/pnas.0500334102.
- [47] S. Salti, F. Tombari, and L. Di Stefano, "SHOT: Unique signatures of histograms for surface and texture description," *Computer Vision and Image Understanding*, vol. 125, pp. 251-264, 2014.
- [48] G. Shamai and R. Kimmel, "Geodesic distance descriptors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 6410-6418.
- [49] M. Ovsjanikov, M. Ben-Chen, J. Solomon, A. Butscher, and L. Guibas, "Functional maps: a flexible representation of maps between shapes," *ACM Transactions on Graphics (ToG)*, vol. 31, no. 4, pp. 1-11, 2012.
- [50] D. Nogneng and M. Ovsjanikov, "Informative descriptor preservation via commutativity for shape matching," in *Computer Graphics Forum*, 2017, vol. 36, no. 2: Wiley Online Library, pp. 259-267.
- [51] E. Rodolà, L. Cosmo, M. M. Bronstein, A. Torsello, and D. Cremers, "Partial functional correspondence," in *Computer graphics forum*, 2017, vol. 36, no. 1: Wiley Online Library, pp. 222-236.
- [52] O. Litany, E. Rodolà, A. M. Bronstein, and M. M. Bronstein, "Fully spectral partial shape matching," in *Computer Graphics Forum*, 2017, vol. 36, no. 2: Wiley Online Library, pp. 247-258.
- [53] J. Ren, A. Poulencard, P. Wonka, and M. Ovsjanikov, "Continuous and orientation-preserving correspondences via functional maps," *ACM Transactions on Graphics (ToG)*, vol. 37, no. 6, pp. 1-16, 2018.
- [54] J.-M. Roufousse, A. Sharma, and M. Ovsjanikov, "Unsupervised deep learning for structured shape matching," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1617-1627.
- [55] O. Halimi, O. Litany, E. R. Rodolà, A. M. Bronstein, and R. Kimmel, "Unsupervised Learning of Dense Shape Correspondence," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 15-20 June 2019 2019, pp. 4365-4374, doi: 10.1109/CVPR.2019.00450.
- [56] E. Rodolà, M. Moeller, and D. Cremers, "Point-wise map recovery and refinement from functional

- correspondence," *arXiv preprint arXiv:1506.05603*, 2015.
- [57] E. Rodola, M. Möller, and D. Cremers, "Regularized pointwise map recovery from functional correspondence," in *Computer Graphics Forum*, 2017, vol. 36, no. 8: Wiley Online Library, pp. 700-711.
- [58] D. Ezuz and M. Ben-Chen, "Deblurring and denoising of maps between shapes," in *Computer Graphics Forum*, 2017, vol. 36, no. 5: Wiley Online Library, pp. 165-174.
- [59] M. Vestner, R. Litman, E. Rodola, A. Bronstein, and D. Cremers, "Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3327-3336.
- [60] S. Melzi, J. Ren, E. Rodola, A. Sharma, P. Wonka, and M. Ovsjanikov, "Zoomout: Spectral upsampling for efficient shape correspondence," *arXiv preprint arXiv:1904.07865*, 2019.
- [61] G. Pai, J. Ren, S. Melzi, P. Wonka, and M. Ovsjanikov, "Fast sinkhorn filters: Using matrix scaling for non-rigid shape correspondence with functional maps," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 384-393.
- [62] O. Litany, T. Remez, E. Rodola, A. Bronstein, and M. Bronstein, "Deep functional maps: Structured prediction for dense shape correspondence," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 5659-5667.
- [63] O. Halimi, O. Litany, E. Rodola, A. M. Bronstein, and R. Kimmel, "Unsupervised learning of dense shape correspondence," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4370-4379.
- [64] N. Donati, E. Corman, and M. Ovsjanikov, "Deep orientation-aware functional maps: Tackling symmetry issues in shape matching," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 742-751.
- [65] D. Cao, P. Roetzer, and F. Bernard, "Unsupervised learning of robust spectral shape matching," *arXiv preprint arXiv:2304.14419*, 2023.
- [66] H. Chen and B. Bhanu, "3D free-form object recognition in range images using local surface patches," *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1252-1262, 2007.
- [67] M. Botsch, L. Kobbelt, M. Pauly, P. Alliez, and B. Lévy, *Polygon mesh processing*. CRC press, 2010.
- [68] R. Jonker and T. Volgenant, "A shortest augmenting path algorithm for dense and sparse linear assignment problems," in *DGOR/NSOR: Papers of the 16th Annual Meeting of DGOR in Cooperation with NSOR/Vorträge der 16. Jahrestagung der DGOR zusammen mit der NSOR*, 1988: Springer, pp. 622-622.
- [69] Y. Li, H. Takehara, T. Taketomi, B. Zheng, and M. Nießner, "4dcomplete: Non-rigid motion estimation beyond the observable surface," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 12706-12716.
- [70] S. Zuffi, A. Kanazawa, D. W. Jacobs, and M. J. Black, "3D menagerie: Modeling the 3D shape and pose of animals," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 6365-6373.
- [71] O. Azencot and R. Lai, "A Data-Driven Approach to Functional Map Construction and Bases Pursuit," in *Computer Graphics Forum*, 2021, vol. 40, no. 5: Wiley Online Library, pp. 97-110.
- [72] V. G. Kim, Y. Lipman, and T. Funkhouser, "Blended intrinsic maps," *ACM transactions on graphics (TOG)*, vol. 30, no. 4, pp. 1-12, 2011.
- [73] D. Morgan, G. P. Subramanian, S.-J. Chung, and F. Y. Hadaegh, "Swarm assignment and trajectory optimization using variable-swarm, distributed auction assignment and sequential convex programming," *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1261-1285, 2016.
- [74] R. Jonker and T. Volgenant, "Improving the Hungarian assignment algorithm," *Operations research letters*, vol. 5, no. 4, pp. 171-175, 1986.

ضمیمه (الف) محاسبه و تحلیل پیچیدگی زمانی الگوریتم ۳ (LAPJV)

در این بخش پیچیدگی محاسباتی الگوریتم ۳ با ارایه شبه کد ++C تحلیل و محاسبه شده است. الگوریتم ۳ یا LAPJV شامل دو فرآیند مقدماتی و بهبود با اصلاحات اضافی است. فرآیند مقدماتی در سه مرحله، کاهش ستون (CRS)، انتقال کاهش (RTS) و کاهش ردیف تقویتی (ARRS) اجرا می‌شود [۶۸]. بنابراین، پیچیدگی زمانی هر دو فرآیند با مراحل مربوطه در نظر گرفته می‌شود. در مرحله کاهش ستون از فرآیند مقدماتی (خطوط ۲۹ تا ۳۸ شبه کد) دو حلقه تو در تو وجود دارد که یکی روی سطرها و دیگری روی ستون‌ها تکرار می‌شوند. هر تکرار شامل عملیات زمان ثابت است، بنابراین پیچیدگی زمانی مرحله کاهش ستون (CRS) برابر با $O(n^2)$ است که n بعد ماتریس هزینه است. در مرحله انتقال کاهش (خطوط ۳۹ تا ۴۹ از شبه کد) نیز دو حلقه تو در تو وجود دارد که هر دو حلقه داخلی و بیرونی n بار و در مجموع منجر به n^2 بار تکرار می‌شود. بنابراین، پیچیدگی زمانی مرحله انتقال کاهش (RTS) برابر با $O(n^2)$ است. در مرحله کاهش ردیف تقویتی (خطوط ۵۰ تا ۷۷ شبه کد) از طریق هر عنصر ماتریس هزینه، تکرارها در دو حلقه تو در تو وجود دارد که باعث می‌شود پیچیدگی زمانی این مرحله از درجه دوم باشد. به عبارت دیگر، مرحله کاهش ردیف تقویتی (ARRS) برابر $O(n^2)$ است. به همین ترتیب، پیچیدگی زمانی فرآیند بهبود با اصلاحات اضافی (خطوط ۷۸ تا ۱۱۸ شبه کد) در الگوریتم ۳ برابر $O(n^3)$ است. در حالت کلی، مجموع پیچیدگی‌های محاسبه شده بیانگر پیچیدگی زمانی الگوریتم ۳ است. نتایج آزمایش‌ها نشان می‌دهند که پیچیدگی زمانی الگوریتم ۳ (LAPJV) بسیار کمتر از الگوریتم‌های مشابه مانند الگوریتم حراج [۷۳]، کوهن-مانکراس (KM) و نسخه بهبود یافته آن [۷۴] است، زیرا بیشتر انتساب‌های نقاط تا پایان فرآیند مقدماتی تکمیل می‌شوند و فقط درصد کمی از نقاط باقی می‌مانند که هنوز تخصیص نیافته‌اند و باید در فرآیند بهبود با اصلاحات اضافی استفاده شوند. در ادامه، دلیل پیچیدگی محاسباتی کم الگوریتم ۳ بررسی شده است. نتایج تخصیص و پیچیدگی زمانی فرآیندهای مقدماتی (شامل زیرمراحل CRS، RTS، ARR) و بهبود برای ۵۰۰۰ نقطه از ۱۰۰ جفت مدل محاسبه شده است. آمار انتساب‌ها در شکل ۱۰ نشان داده شده‌اند. میانگین درصد نقاط تخصیص یافته در فرآیندهای مقدماتی و بهبود به ترتیب، ۹۹/۹۹۸۷۴٪ و ۰/۰۰۱۲۶٪ محاسبه شده است. به عبارت دیگر، نتایج نشان می‌دهند که بیش از ۹۹/۹۹ درصد نقاط در فرآیند مقدماتی با پیچیدگی $O(n^2)$ انتساب داده می‌شوند. بنابراین، فقط ۰/۰۰۱۲۶ درصد از نقاط مدل که هنوز تخصیص نیافته‌اند در فرآیند بهبود با پیچیدگی $O(n^3)$ پردازش می‌شوند. بنابراین، فشرده‌ترین بخش الگوریتم LAPJV به فرآیند مقدماتی مربوط می‌شود. به طور میانگین، پیچیدگی محاسباتی الگوریتم LAPJV برابر $O(n^2)$ است.



شکل ۱۰. مقایسه آماری تعداد انتساب نقاط برای ۵۰۰۰ نقطه (رأس). (الف) آمار تعداد انتساب‌ها بعد از مرحله کاهش ستون (CRS). (ب) آمار تعداد انتساب‌ها بعد از مرحله کاهش ستون (RTS). (پ) آمار تعداد انتساب‌ها بعد از مرحله کاهش ستون (ARRS). (ج) آمار تعداد انتساب‌ها بعد از فرآیند بهبود با اصلاحات اضافی.

```
1. #if !defined TRUE
2. #define TRUE 1
3. #endif
4. #if !defined FALSE
5. #define FALSE 0
6. #endif
7. /***** DATA TYPES *****/
8. typedef int boolean;
10. #define B 1e5
11. #define resolution 1e-4
12. /***** TYPES *****/
13. typedef int Row;
14. typedef int Col;
15. typedef double Cost;
16. Cost lap(int n, Cost **costofassign, Col *RowSol, Row *ColSol, Cost *u, Cost *v)
17. {
18. boolean unassign;
19. Row i, imin, numf = 0, prvnumf, f, i0, k, fRow, *pred, *fr;
20. Col j, j1, j2, endofp, last, low, up, *Collist, *match;
21. Cost min, h, umin, usubmin, v2, *d;
22. fr = new Row[n]; // لیستی از ردیف‌های تخصیص نیافته
23. Collist = new Col[n]; // لیستی از ستون‌های اسکن شده تخصیص نیافته
24. match = new Col[n]; // شمارش تعداد دفعاتی که یک ردیف تخصیص یافته
25. d = new Cost[n]; // متغیر فاصله برای محاسبه مسیر در فاز تقویت
26. pred = new Row[n]; // متغیر ردیف شمار برای ستون در فاز تقویت
    تعیین اینکه چند بار یک ردیف در مرحله کاهش ستون انتساب داده می‌شود.
27. for (i = 0; i < n; i++)
28. match[i] = 0;
    مرحله کاهش ستون
29. for (j = n-1; j >= 0; j--)
30. { min = costofassign[0][j]; imin = 0;
31. for (i = 1; i < n; i++)
32. if (costofassign[i][j] < min)
33. { min = costofassign[i][j]; imin = i;}
34. v[j] = min;
35. if (++match[imin] == 1)
36. { Rowsol[imin] = j; ColSol[j] = imin;}
```

```
37. else
38. Colsol[j] = -1;}
```

مرحله انتقال کاهش

```
39. for (i = 0; i < n; i++)
40. if (match[i] == 0)
41. fr[numf++] = i;
42. else
43. if (match[i] == 1)
44. { j1 = Rowsol[i]; min = B;
45. for (j = 0; j < n; j++)
46. if (j != j1)
47. if (costofassign[i][j] - v[j] < min)
48. min = costofassign[i][j] - v[j]; v[j1] = v[j1] - min;
49. }
```

مرحله کاهش ردیف تقویتی

```
50. int loopcnt = 0;
51. do
52. { loopcnt++; k = 0; prvnumf = numf; numf = 0;
53. while (k < prvnumf)
54. { i = fr[k]; k++; umin = costofassign[i][0] - v[0]; j1 = 0; usubmin = B;
55. for (j = 1; j < n; j++)
56. { h = costofassign[i][j] - v[j];
57. if (h < usubmin)
58. if (h >= umin)
59. { usubmin = h; j2 = j; }
60. else
61. { usubmin = umin; umin = h; j2 = j1; j1 = j; }
62. }
63. i0 = Colsol[j1];
64. if (usubmin - umin > resolution)
65. v[j1] = v[j1] - (usubmin - umin);
66. else
67. if (i0 >= 0)
68. { j1 = j2; i0 = Colsol[j2]; }
69. Rowsol[i] = j1; Colsol[j1] = i;
70. if (i0 >= 0)
71. if (usubmin - umin > resolution)
72. fr[--k] = i0;
73. else
74. fr[numf++] = i0;
75. }
76. }
77. while (loopcnt < 2);
```

فاز بهبود با اصلاحات اضافی برای ردیف‌های تخصیص نیافته

```
78. for (f = 0; f < numf; f++)
79. { fRow = fr[f];
80. for (j = 0; j < n; j++)
81. { d[j] = costofassign[fRow][j] - v[j]; pred[j] = fRow; Collist[j] = j; }
82. low = 0; up = 0; unassign = FALSE;
83. do
```

```

84. {
85.   if (up == low)
86.     { last = low - 1; min = d[Collist[up++]];
87.       for (k = up; k < n; k++)
88.         { j = Collist[k]; h = d[j];
89.           if (h <= min)
90.             {
91.               if (h < min)
92.                 { up = low; min = h; }
93.               Collist[k] = Collist[up]; Collist[up++] = j; }
94.             }
95.           for (k = low; k < up; k++)
96.             if (Colsol[Collist[k]] < 0)
97.               { endofp = Collist[k]; unassign = TRUE; break; }
98.           }
99.         if (!unassign)
100.        {
    
```

به روز رسانی فاصله بین متغیر فاصله (fRow) و تمام ستون‌های اسکن نشده توسط ستون اسکن شده بعدی

```

101.    j1 = Collist[low]; low++; i = Colsol[j1]; h = costofassign[i][j1] - v[j1] - min;
102.    for (k = up; k < n; k++)
103.      {
104.        j = Collist[k]; v2 = costofassign[i][j] - v[j] - h;
105.        if (v2 < d[j])
106.          {
107.            pred[j] = i;
108.            if (v2 == min)
109.              if (Colsol[j] < 0)
110.                { endofp = j; unassign = TRUE; break; }
111.            else
112.              { Collist[k] = Collist[up]; Collist[up++] = j; }
113.            d[j] = v2;
114.          }
115.      }
116.    }
117.  }
118.  while (!unassign);
    
```

به روز رسانی ارزش ستون‌ها

```

119.  for (k = 0; k <= last; k++)
120.    {
121.      j1 = Collist[k]; v[j1] = v[j1] + d[j1] - min; }
122.  do
123.    { i = pred[endofp]; Colsol[endofp] = i; j1 = endofp; endofp = Rowsol[i]; Rowsol[i] = j1; }
124.  while (i != fRow);
125. }
    
```

محاسبه هزینه بهینه

```

126. Cost lapCost = 0;
127. for (i = 0; i < n; i++)
128. { j = Rowsol[i]; u[i] = costofassign[i][j] - v[j]; lapCost = lapCost + costofassign[i][j]; }
129. delete[] pred; delete[] fr; delete[] Collist; delete[] match; delete[] d;
130. return lapCost; }
    
```

Detection of Abnormal Driver Behaviors Using Local Binary Patterns and Histogram of Oriented Gradients

Nima Kaveh Maryan, Hossein Ebrahimnezhad*, Ahmadreza Yeganeh Mehr

Computer Vision Res. Lab., Faculty of Electrical and Computer Engineering,
Sahand University of Technology, Tabriz, Iran
E-mails: kaveh96@sut.ac.ir; ebrahimnezhad@sut.ac.ir; ar_yeganehmehr98@sut.ac.ir



Sahand University
of Technology

DOR:

[20.1001.1.23223146.1404.12.1.2.5](https://doi.org/10.1001.1.23223146.1404.12.1.2.5)

Journal of Nonlinear
Systems in Electrical
Engineering

Vol.12, No.1

Spring and Summer 2025

ISSN: 2322 – 3146

<http://jnsee.sut.ac.ir>

Abstract

Nowadays, abnormal and unconventional driver behaviors have become a serious challenge in society, necessitating fundamental solutions and preventive programs. Lack of vehicle control skills, fatigue, drowsiness, and distraction are among the most common causes of traffic accidents. In computer vision, detecting abnormal driver behaviors is both crucial and challenging. Therefore, our goal is to identify abnormal driver behaviors using computer vision techniques, including driver drowsiness, smoking, and mobile phone usage while driving. In this study, we employ two feature extraction methods: Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG) to detect abnormal driver behaviors. The first step involves detecting the driver's face, which is identified using 68 key facial landmarks. The face is then divided into four regions: eyes, mouth, right ear, and left ear. In the first method, feature extraction is performed using the LBP algorithm for each behavior, followed by classification using a Support Vector Machine (SVM). In the second method, feature extraction is conducted using HOG, and the extracted features are classified with an SVM. The results indicate that the HOG-based feature extraction method improves classification accuracy in the SVM model, achieving better performance compared to LBP features.

Keywords

Drowsiness; Driver Safety; Machine Learning; Local Binary Patterns; Histogram of Oriented Gradients

1. Short Introduction

Driver distraction detection is a crucial aspect of advanced driver assistance systems aimed at improving transportation safety. Research has concentrated on analyzing driver eye behavior and head movements, utilizing systems like PERCLOS to measure eye closure. Recent advancements include the application of Convolutional Neural Networks (CNNs) and deep learning models such as ResNet and MobileNetV2 to identify distraction patterns. Additionally, studies have tackled challenges in monitoring driver behavior, including visibility obstructions and lighting conditions. Computer vision-based methods have been proposed to effectively detect drowsiness and distraction in drivers.

2. Proposed Work

The proposed method for detecting the driver's status and state involves several steps. First, face detection is performed as the main preprocessing step. After detecting the face, key facial points are extracted, marking the boundaries of facial components such as the eyes, eyebrows, nose, mouth, and jaw. Mapping these points onto the driver's face image allows for precise segmentation of various facial regions. To assess the driver's facial condition and analyze features related to fatigue and distraction, detailed segmentation of areas including the left eye, right eye, left eyebrow, right eyebrow, nose, mouth, left ear, right ear, and jaw is conducted.

Once the facial regions are separated, image features need to be extracted for analyzing and classifying driver behavior. In this phase, two methods, Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP), are used for feature extraction. After extracting features from each facial region, these features are classified using a Support Vector Machine (SVM) algorithm. With the extracted features and the SVM model, different driver states are identified to determine if the driver is experiencing fatigue or distraction.

To evaluate the effectiveness of the proposed method, classification accuracy results for the two feature extraction methods, HOG + SVM and LBP + SVM, are compared. The average accuracy of 88.5% for the HOG feature extraction method with SVM classification demonstrates superior performance due to its ability to extract structural facial features for detecting driver fatigue and distraction. Overall, HOG + SVM performs significantly better than LBP + SVM, which achieves average accuracy of 61.5%, indicating lower efficiency in detecting various facial states.

Although the proposed method relies on manually extracted features like HOG + SVM, it has limitations compared to deep learning-based approaches. Deep networks can learn abstract and complex features, showing better performance under varying lighting conditions, different angles, and random driver movements through techniques such as data augmentation and adaptive learning. Furthermore, hybrid models with advanced settings like multi-objective optimization and adaptive learning enhance detection accuracy. For instance, CNN-LSTM models can analyze temporal changes in driver behavior, while the HOG + SVM method only analyzes static facial features.

3. Conclusion

In conclusion, this study addresses the critical challenge of detecting driver fatigue and distraction to enhance road safety. Our proposed image processing-based system, which includes face detection, facial component segmentation, and driver status classification, achieved an accuracy of 88.5% with the HOG + SVM method, outperforming LBP + SVM at 61.5%. While HOG + SVM is a fast and cost-effective solution, it has limitations under severe environmental changes compared to deep learning approaches. Advanced models like ViT and CNN + LSTM offer higher accuracy but are less suitable for embedded systems due to their high computational requirements. Future research should explore hybrid models that integrate traditional feature extraction with neural networks to improve the accuracy and robustness of driver distraction detection systems.

تشخیص رفتارهای غیر طبیعی رانندگان از طریق آموزش الگوهای باینری محلی و هیستوگرام گرادیان جهت دار

نیما کاوه مریان^۱، حسین ابراهیم نژاد^{۲*}، احمد رضا یگانه مهر^۳

^۱ کارشناسی ارشد مهندسی برق- مخابرات، آزمایشگاه تحقیقاتی بینایی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی تبریز (سهند)، n_kaveh96@sut.ac.ir

^۲ نویسنده مسؤؤل، استاد گروه برق- مخابرات، آزمایشگاه تحقیقاتی بینایی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی تبریز (سهند)، ebrahimnezhad@sut.ac.ir

^۳ کارشناسی ارشد مهندسی برق- مخابرات، آزمایشگاه تحقیقاتی بینایی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی تبریز (سهند)، ar_yeganehmehr98@sut.ac.ir



دانشگاه صنعتی سهند

DOR:

[20.1001.1.23223146.1404.12.1.2.5](https://doi.org/10.1001.1.23223146.1404.12.1.2.5)

نشریه سالانه علمی-فنی در مهندسی برق

دوره ۱۲ - شماره ۱

بهار و تابستان ۱۴۰۴

صفحات ۳۶ الی ۵۶

ISSN: 2322-3146

<http://jnsee.sut.ac.ir>

تاریخ پذیرش: ۱۴۰۳/۱۲/۲۲

تاریخ بازنگری: ۱۴۰۳/۶/۲۰

تاریخ ارسال: ۱۴۰۳/۴/۲۷

چکیده

واژه‌های کلیدی

امروزه رفتارهای غیرطبیعی و نامتعارف رانندگان به یک چالش جدی در جامعه تبدیل شده است که ضروری است برای پیشگیری از آن، راهکارها و برنامه‌های اساسی در نظر گرفته شود. کمبود مهارت در کنترل خودرو، خستگی، خواب آلودگی و حواس پرتی از جمله بیشترین دلایل وقوع تصادفات در کشور است. در بینایی کامپیوتر، تشخیص رفتارهای غیرطبیعی رانندگان بسیار مهم و چالش برانگیز می‌باشد. براین اساس هدف ما تشخیص رفتارهای غیرطبیعی رانندگان با استفاده از بینایی کامپیوتر است که شامل خواب آلودگی راننده، استعمال دخانیات و صحبت کردن با تلفن همراه می‌باشد. در این تحقیق، سعی شده از دو روش آموزش ویژگی‌های باینری محلی (LBP) و هیستوگرام گرادیان جهت‌دار (HOG) برای تشخیص رفتارهای غیرطبیعی رانندگان استفاده شود. اولین گام تشخیص چهره فرد مورد نظر می‌باشد. برای این امر، چهره فرد با ۶۸ نقطه حیاتی تشخیص داده و در ادامه چهره را به چهار بخش چشم، دهان، گوش راست و گوش چپ تقسیم بندی شده است. در روش اول برای هر کدام از وضعیت‌ها با استفاده از الگوریتم الگوهای باینری محلی استخراج ویژگی انجام شده و در ادامه با استفاده از SVM طبقه‌بندی صورت گرفت. همچنین در روش دوم با هیستوگرام گرادیان جهت‌دار استخراج ویژگی گردیده و ویژگی‌های مربوطه به طبقه‌بندی SVM اعمال شده است. نتایج نشان می‌دهد که ویژگی هیستوگرام گرادیان جهت‌دار مورد استفاده شده در این کار دقت ارزیابی را در طبقه‌بندی SVM بیشتر کرده است و به نتایج بهتری نسبت به ویژگی الگوهای باینری محلی دست یافته است.

خواب آلودگی،

حواس پرتی،

استخراج ویژگی،

ایمنی راننده،

یادگیری ماشین،

الگوهای باینری محلی،

هیستوگرام گرادیان جهت‌دار.

۱- مقدمه

سالانه تعداد زیادی تصادف در سراسر جهان اتفاق می افتد و این امر باعث از دست رفتن جان بسیاری از افراد می شود. خواب آلودگی و حواس پرتی یکی از علل های اصلی تصادفات منجر به مرگ بشمار می آیند، که طبق گزارش سازمان امنیت ترافیک بزرگراه ملی^۱ (NHTSA) در آمریکا، در سال ۲۰۲۰ حواس پرتی جان ۳۱۴۲ نفر و همچنین خواب آلودگی جان ۶۹۷ نفر را گرفته است [۱]. حواس پرتی باعث کاهش آگاهی و تصمیم گیری رانندگان شده در نتیجه پاسخ به رویدادهای رانندگی با تاخیر و اختلال در سرعت همراه می شود. رانندگان حواس پرت، کنترل وسیله نقلیه خود را از دست داده و با وسایل نقلیه یا اشیاء ثابت تصادف می کنند. اداره ملی ایمنی ترافیک بزرگراه طی گزارشی سه رفتار رانندگان، شامل حرف زدن با تلفن همراه، ارسال پیامک و استعمال دخانیات را که باعث حواس پرتی رانندگان برای مدت طولانی از جاده می شود، عامل تصادفات گزارش کرده است [۲]. انجام این رفتار در حین رانندگی بر عملکرد رانندگان اثر می گذارد. به همین دلیل آن ها زمان کمتری را صرف بررسی آینه ها، چراغ راهنمایی و تقاطع ها می کنند. تصادفات وسایل نقلیه موتوری به دلیل حواس پرتی رانندگی به طور پیوسته در سال های اخیر افزایش یافته است و محققان، کارشناسان را بر آن داشت تا برای کاهش حوادث و رانندگی ایمن تحقیقات گسترده ای در سال های گذشته تا به امروز انجام دهند. پیشرفت های سریع در رایانه ها و فناوری های بی سیم، دستگاه های پیچیده و متنوعی را ساخته است که افراد می توانند در حین رانندگی از آنها استفاده کنند. با این حال، استفاده از چنین دستگاه هایی در حین رانندگی هم می تواند حواس راننده را از تمرکز بر رانندگی منحرف کند و هم می تواند در خودروهای هوشمند در مواقع حواس پرتی راننده، هشدارهایی مبنی بر خطر و عدم توجه راننده به او نشان دهد و در صورت عدم توجه به هشدارها، کنترل خودرو را به حالت خودکار در بیاورد.

روش های تشخیص حواس پرتی راننده را می توان بر اساس نوع داده آنها دسته بندی کرد [۳]. منابع داده برای تشخیص حواس پرتی راننده عبارتند از: (۱) نشانگرهای مبتنی بر رفتار خودرو یا رانندگی، (۲) فیزیولوژیکی: سیگنال های بیولوژیکی راننده، (۳) متغیرهای رفتاری (رفتارهای بصری رانندگی) یا متغیرهای فیزیکی راننده براساس ویژگی های بصری، مانند تصاویر یا فیلم های گرفته شده از راننده و (۴) اقدامات ترکیبی که برای شناسایی حالت راننده برای کمک به هدایت ایمن خودرو و عدم بروز تصادف ضروری می باشد. به همین دلیل سیستمی که بتواند خواب آلودگی و حواس پرتی راننده را تشخیص دهد، از اهمیت بالایی برخوردار می باشد. روش های ارائه شده در تشخیص خواب آلودگی و حواس پرتی راننده به دو دسته بصری و غیربصری تقسیم بندی می شوند. در چند دهه گذشته محققان زیادی در زمینه روش های غیربصری تحقیق نموده اند. در روش غیر بصری از سیگنال هایی همچون EEG^2 ، EMG^3 ، EOG^4 و ECG^5 استخراج ویژگی می شود [۴]، که دقت بالایی دارد اما برای به دست آوردن این ویژگی ها باید الکترودهایی به بدن راننده متصل شود که این امر برای فرد خوشایند نیست. در روش بصری رفتار راننده از جمله صحبت کردن با تلفن همراه، استعمال دخانیات و خواب آلودگی راننده مورد نظارت قرار می گیرد و سپس با استفاده از پردازش تصویر و بینایی کامپیوتر، ویژگی های لازم استخراج می شوند. استخراج ویژگی از تصاویر شامل روش های متعددی است که بسته به کاربرد موردنظر می توان از آن ها بهره برد. در پردازش تصویر کلاسیک، تکنیک هایی مانند، تبدیل ویژگی مستقل از مقیاس (SIFT)^۶ و (SURF)^۷ به منظور شناسایی نقاط کلیدی و توصیفگرهای مقاوم در برابر تغییرات مقیاس و چرخش استفاده می شوند. همچنین، ویژگی های مبتنی بر گرادیان نظیر هیستوگرام گرادیان جهت دار (HOG)^۸ برای توصیف ساختارهای لبه ای و تشخیص اشیا کاربرد دارد. یکی دیگر از

¹ National Highway Traffic Safety Administration (NHTSA)

² Electroencephalography

³ Electromyography

⁴ Electrooculography

⁵ Electrocardiography

⁶ Scale-Invariant Feature Transform

⁷ Speeded-Up Robust Features

⁸ Histogram of Oriented Gradients

روش‌های مهم، الگوی باینری محلی (LBP)^۱ است که با بررسی الگوهای بافتی بر اساس مقایسه شدت روشنایی پیکسل‌های مجاور، ویژگی‌های مقاومی را در برابر تغییرات روشنایی و چرخش استخراج می‌کند و در کاربردهایی مانند تشخیص چهره و دسته‌بندی تصاویر بافتی مورد استفاده قرار می‌گیرد. در روش‌های مدرن‌تر، شبکه‌های عصبی پیچشی (CNN) به‌عنوان استخراج‌کننده‌های خودکار ویژگی، توانایی یادگیری نمایش‌های انتزاعی و سلسله‌مراتبی از تصویر را دارند که در کاربردهایی مانند دسته‌بندی، تشخیص اشیا و بازیابی تصویر عملکرد برتری ارائه می‌دهند. در حوزه بینایی ماشین و ردیابی، Optical Flow به تحلیل حرکت پیکسل‌ها در دنباله‌ای از فریم‌ها کمک می‌کند و در کاربردهایی مانند تشخیص فعالیت و تحلیل ویدئو به کار می‌رود. همچنین، مدل‌های مبتنی بر تبدیل‌ها مانند Wavelet Transform برای استخراج ویژگی‌های فرکانسی-مکانی استفاده می‌شوند که در پردازش تصاویر پزشکی و فشرده‌سازی داده‌ها مفید هستند. در نهایت، روش‌های یادگیری عمیق نظیر ViT^۲، به‌ویژه در پردازش تصاویر پیچیده، توانسته‌اند با یادگیری وابستگی‌های طولانی‌مدت در تصویر، عملکرد قابل توجهی در وظایفی مانند طبقه‌بندی کهکشان‌ها و تشخیص الگوهای پیچیده ارائه دهند. ما در این پژوهش از ۲ روش استخراج ویژگی شامل، (۱) هیستوگرام گرادیان جهت‌دار (HOG)، (۲) الگوی باینری محلی (LBP)، استفاده نموده‌ایم و با ارائه راهکاری، خواب‌آلودگی و عوامل منجر به حواس‌پرتی رانندگان را با بررسی رفتارهای چشم، صحبت با تلفن و کشیدن سیگار تشخیص می‌دهیم. ساختار تحقیق به این صورت است، که بخش اول شامل مقدمه و مفاهیم اولیه در مورد تشخیص حواس‌پرتی و خستگی رانندگان می‌باشد. بخش دوم، به مرور مختصری بر کارهای انجام شده در مورد رفتارهای غیرنرمال رانندگان پرداخته است. در بخش سوم، روش پیشنهادی مورد بحث قرار خواهد گرفت. در بخش چهارم، نتایج به‌دست آمده از روش انجام شده مورد ارزیابی قرار می‌گیرد و در بخش پنجم، نتیجه‌گیری کلی از تحقیق انجام می‌گیرد.

۲- کارهای مربوطه

تشخیص حواس‌پرتی راننده یکی از حوزه‌های مهم در سیستم‌های پیشرفته کمک‌راننده (ADAS) و افزایش ایمنی حمل‌ونقل است که تاکنون مطالعات متعددی در این زمینه انجام شده است. این پژوهش‌ها عمدتاً بر تحلیل رفتار چشمی و حرکات سر، استفاده از یادگیری عمیق برای پردازش تصاویر، پردازش سیگنال‌های بیولوژیکی و ترکیب چندمنظوره داده‌ها متمرکز بوده‌اند. یکی از مهم‌ترین روش‌های تشخیص حواس‌پرتی، تحلیل رفتار چشمی راننده است. مطالعه‌ای توسط جای و همکاران، سیستم PERCLOS را معرفی کرد که درصد بسته بودن چشم‌ها را اندازه‌گیری کرده و به‌عنوان شاخصی برای تشخیص خواب‌آلودگی و حواس‌پرتی به کار گرفته شد [۵]. مطالعه دیگری توسط ویستنه و همکاران، نشان داد که تحلیل مسیر نگاه راننده و ارتباط آن با محیط جاده‌ای می‌تواند نقش مهمی در شناسایی عدم تمرکز داشته باشد [۶]. روش‌های مبتنی بر یادگیری ماشین مانند SVM و Random Forest نیز برای تحلیل الگوهای چشم و سر راننده به کار گرفته شده‌اند [۷]. با پیشرفت بینایی کامپیوتر، مطالعاتی از شبکه‌های عصبی پیچشی (CNN) و مدل‌های یادگیری عمیق برای تشخیص الگوهای پیچیده چهره و حرکات سر بهره برده‌اند. پژوهش ابولناگا و همکاران، مدلی مبتنی بر ResNet را ارائه داد که ویژگی‌های صورت راننده را برای تشخیص توجه وی به مسیر استخراج می‌کرد [۸]. در یک مطالعه دیگر، باهتی و همکاران، از مدل MobileNetV2 برای تشخیص استفاده از تلفن همراه حین رانندگی استفاده کردند و نشان دادند که این روش عملکرد بالایی در مقایسه با مدل‌های کلاسیک دارد [۹]. علاوه بر این، اخیراً استفاده از ویژن ترانسفورمر (ViT) در تحلیل تصاویر رانندگان مورد بررسی قرار گرفته است که دقت بالاتری در تشخیص الگوهای غیرعادی نشان داده است [۱۰]. برخی مطالعات از سیگنال‌های EEG و ECG برای شناسایی وضعیت ذهنی راننده استفاده کرده‌اند. پژوهش جانگ و همکاران، نشان داد که

¹ Local Binary Patterns

² Vision Transformer

سیگنال‌های EEG می‌توانند برای شناسایی کاهش هوشیاری و حواس‌پرتی مورد استفاده قرار گیرند [۱۱]. همچنین، مطالعه‌ای توسط لین و همکاران، به بررسی تغییرات ضربان قلب (HRV) پرداخت و نشان داد که کاهش تمرکز راننده با تغییر در الگوی HRV قابل تشخیص است [۱۲، ۱۳، ۱۴]. محققان برای تعیین میزان خستگی و خواب‌آلودگی راننده با بررسی و استخراج ویژگی در زمان‌های هوشیاری و خواب‌آلودگی بر سیگنال‌های الکتروانسفالوگرام (EEG) و الکتروکاردیوگرام (ECG) توانستند سطح هوشیاری فرد را تعیین کنند و در نتیجه حالت خواب‌آلودگی را در مراحل اولیه تشخیص بدهند. سیگنال EEG یک وسیله فیزیولوژیکی غیرتهاجمی برای اندازه‌گیری فعالیت مغز است و نزدیک‌ترین رابطه را با خواب‌آلودگی دارد [۱۵]. به‌ویژه، در شرایط مشابه، EEG بهتر از هر سیگنال فیزیولوژیکی دیگر عمل می‌کند [۱۶]. سیگنال EEG دارای باندهای فرکانسی متفاوتی است. بررسی این سیگنال در هنگام خواب‌آلودگی نشان داده است که باندهای آلفا و تتا تغییرات گسترده‌ای نسبت به سایر باندهای فرکانسی دارند. همچنین افزایش تغییرات قدرت در باند فرکانسی تتا و کاهش آن در باند فرکانسی آلفا، یکی از نشانه‌های خواب‌آلودگی محسوب می‌شود [۱۷]. در روش بصری با استفاده از بینایی کامپیوتر، حرکات رانندگان مورد نظارت قرار می‌گیرد. برای این امر نیاز به نصب دوربین در داخل وسیله نقلیه برای بررسی رفتارهای رانندگان وجود دارد. این رفتارها شامل حرکات صورت مانند پلک زدن مداوم، چشمان باز و بسته، خمیازه کشیدن، تکان دادن یا تکان ندادن سر است [۱۸]. البته در این رویکرد می‌توان به چالش‌هایی همچون پوشاندن چشم به وسیله عینک آفتابی یا عینک معمولی، پوشاندن دهان توسط ماسک یا دست، زاویه دید راننده، شرایط نوری، حالت موی راننده و همچنین داشتن نقاب یا کلاه اشاره کرد [۱۹]. در این روش برای نظارت بر رانندگان می‌توان از رفتارهای مختلف برای تشخیص خواب‌آلودگی و حواس‌پرتی که یکی از علل اصلی تصادفات مرگبار است، استفاده کرد. استخراج ویژگی از چشم‌ها و رفتار رانندگان مانند صحبت کردن با تلفن همراه و استعمال دخانیات از جمله این روش‌ها است. در پژوهشی، سلیمان و همکاران [۲۰] از یک روش مبتنی بر یادگیری ماشین برای تشخیص خواب‌آلودگی استفاده کردند. در این روش، دوربینی متصل به داشبورد تصاویری از چهره راننده اخذ کرده، سپس با الگوریتم Viola-Jones و ویژگی‌های Haar چهره را تشخیص داده و ویژگی‌های چشم و دهان را استخراج کردند. در ادامه از روش درخت تصمیم‌گیری برای طبقه‌بندی حالت باز و بسته چشم استفاده شد. مبین عاریف و همکاران [۲۱] از روش پردازش تصویر زمان واقعی برای تشخیص خواب‌آلودگی راننده و سیستم هشدار مبتنی بر نرخ پلک زدن استفاده کردند. عبدالمالک مجاهد و همکاران [۲۲] یک سیستم مانیتورینگ چهره را بر اساس یک توصیفگر بافت چهره فشرده ارائه کردند که می‌تواند متمایزترین ویژگی‌های خواب‌آلودگی را پوشش دهد. در اینجا، چشم‌های هر چهره با استفاده از الگوریتم گروه درختان رگرسیون (ERT) شناسایی شدند که یک الگوریتم قوی و بسیار کارآمد برای محلی‌سازی نشانه‌های چهره محسوب می‌شود. سپس نمایش چندسطحی PML برای استخراج ویژگی‌های محلی هر بلوک در هر سطح با استفاده از توصیفگرهای HOG، COV و LBP انجام شد. داده‌ها با تجزیه و تحلیل مؤلفه اصلی (PCA) کاهش یافتند و در نهایت، طبقه‌بندی با استفاده از ماشین بردار پشتیبان (SVM) انجام شد. ریموند پتوچا و همکاران [۲۳] با استفاده از بینایی کامپیوتر، تلاش کردند رانندگان حواس‌پرت را در مجموعه داده‌ای شامل هفت رفتار مختلف حواس‌پرتی شناسایی کنند. در این تحقیق، ویژگی‌های بینایی کامپیوتر شامل ترکیبی از هیستوگرام گرادبان‌های جهت‌دار (HOG) و توصیفگرهای SIFT همراه با طبقه‌بندی SVM مورد استفاده قرار گرفت. کیم [۲۴] روشی مبتنی بر الگوی باینری محلی (LBP) و SVM برای تشخیص دود در دوربین‌های نظارتی ارائه کرد. شائوچی و همکاران [۲۵] نیز روشی مبتنی بر یادگیری عمیق برای تشخیص فعالیت‌های منجر به حواس‌پرتی راننده ارائه دادند که در آن، تصاویر راننده شامل صحبت کردن با تلفن و پیامک زدن توسط دوربینی متصل به داشبورد وسیله نقلیه جمع‌آوری شد. سیستم مورد نظر از شبکه ResNet-50 که یک CNN پیشرفته است، برای تشخیص حواس‌پرتی رانندگان بهره برد. بررسی مطالعات گذشته نشان می‌دهد که روش‌های مبتنی بر پردازش تصویر و یادگیری عمیق عملکرد بالایی در تشخیص حواس‌پرتی راننده دارند، اما ترکیب این روش‌ها با داده‌های سیگنال‌های زیستی و حسگرهای داخل خودرو می‌تواند دقت سیستم‌های تشخیص را افزایش دهد. به‌ویژه، استفاده از شبکه‌های

عصبی عمیق و مدل‌های مبتنی بر ویژن ترنسفورمر (ViT) رویکردی نوین در این حوزه محسوب می‌شود که می‌تواند در آینده به بهبود سیستم‌های پیشرفته کمک‌رساننده منجر شود.

۳- روش پیشنهادی

ما در این بخش به توضیح روش پیشنهادی خواهیم پرداخت. در این روش تلاش شده است با تمرکز بر استخراج دقیق ویژگی‌های چهره و ترکیب تکنیک‌های مؤثر پردازش تصویر به بهبود دقت سیستم‌های پایش رانندگان کمک شود. پیش از پرداختن به نحوه پیاده سازی روش، لازم است به توضیحاتی در مورد روش‌ها و الگوریتم‌های مورد استفاده پرداخته شود، تا ضمن آشنایی با آنها، بیان روش پیشنهادی راحتتر گردد و از ابهامات در این خصوص جلوگیری شود. در ادامه توضیحاتی درخصوص الگوی باینری محلی (LBP)، هیستوگرام گرادیان جهت دار (HOG) و ماشین بردار پشتیبان (SVM) داده شده است و سپس به تشریح روش به طور گسترده پرداخته شده است.

۳-۱- الگوی باینری محلی (LBP)

الگوهای باینری محلی، یا به اختصار LBP ها، برای اولین بار توسط تیمو اوچالا و همکارانش معرفی شد [۲۶]. برخلاف ویژگی‌های بافت Haralick که یک نمایش کلی از بافت را بر اساس ماتریس هم‌وقوع سطح خاکستری محاسبه می‌کنند، الگوهای باینری محلی (LBP) در عوض نمایش محلی بافت را محاسبه می‌کنند. این نمایش محلی با مقایسه هر پیکسل با همسایگی مقادیر پیکسل اطراف آن انجام می‌شود [۲۷، ۲۸]. اولین مرحله در ساخت توصیفگر LBP، تبدیل تصویر به مقیاس خاکستری است. برای هر پیکسل در تصویر خاکستری، همسایگی به اندازه r اطراف پیکسل مرکزی انتخاب می‌شود. به این صورت که تصویر به قسمت‌های یکسان تقسیم گردیده و برای هر قسمت و تمامی پیکسل‌های موجود با توجه به اختلاف پیکسل‌های همسایه با پیکسل مرکزی یک ماتریس ایجاد می‌گردد. همانطور که بیان شد پیکسل مرکزی مورد نظر در برابر پیکسل‌های همسایه آن مورد بررسی قرار می‌گیرد. اگر شدت پیکسل مرکزی بیشتر از همسایه‌اش باشد، مقدار آن را ۱ و اگر کوچکتر باشد عدد صفر جای گذاری می‌گردد. سپس به صورت پاد ساعت گرد اعداد هر ماتریس که به صورت باینری می‌باشد، خوانده می‌شود. در ادامه عدد دهدهی مربوط به عدد باینری بدست می‌آید. برای همه پیکسل‌ها این کار انجام می‌گردد تا در نهایت همه اعداد دهدهی که از m ناحیه بدست آمده‌اند، یک ماتریس هیستوگرام $m \times n$ را برای هر تصویر به وجود می‌آورند.

۳-۲- هیستوگرام گرادیان جهت دار (HOG)

هیستوگرام گرادیان جهت دار (HOG) یک توصیفگر ویژگی است که توسط ناونیت دالال و بیل تریگر مطرح شده است [۲۹]، اغلب برای استخراج ویژگی‌ها از داده‌های تصویر استفاده می‌شود. این توصیفگر بر این ایده استوار است که ظاهر و شکل یک شی در یک تصویر را می‌توان با استفاده از توزیع گرادیان‌های شدت یا جهت‌های لبه‌ها در نواحی محلی تصویر توصیف کرد. هیستوگرام گرادیان جهت دار (HOG) برخلاف توصیفگر ویژگی‌های لبه، قادر است جهت لبه را نیز ارائه دهد. این کار با استخراج گرادیان جهت‌دار در ناحیه‌های مختلف تصویر انجام شده است. برای بدست آوردن ماتریس هیستوگرام به این صورت عمل می‌شود که ابتدا تصاویر به قطعات کوچک متصل به هم تقسیم می‌شود که به آن‌ها سلول گفته می‌شود. سپس برای هر پیکسل درون یک سلول، گرادیان شدت یا جهت لبه محاسبه می‌شود. این اطلاعات برای ساخت یک هیستوگرام از جهت‌های گرادیان در هر سلول استفاده

می‌شود. در نهایت، توصیفگر HOG با ترکیب هیستوگرام‌های همه سلول‌ها ایجاد می‌شود. برای بهبود دقت و کارایی، هیستوگرام‌های محلی سلول‌ها را می‌توان با در نظر گرفتن یک ناحیه بزرگ‌تر در تصویر به نام بلوک و محاسبه یک مقیاس شدت کلی برای آن بلوک نرمال‌سازی کرد. سپس از این مقیاس برای نرمال کردن مقادیر هیستوگرام در سلول‌های داخل آن بلوک استفاده می‌شود. این فرآیند نرمال‌سازی کنتراست، باعث می‌شود توصیفگر HOG نسبت به تغییرات روشنایی و سایه در تصویر مقاوم‌تر شود. این اطلاعات برای بازنمایی و تشخیص اشیاء در تصاویر مفید هستند و در کاربردهای مختلفی مانند تشخیص عابر پیاده، تشخیص چهره و دسته‌بندی تصاویر استفاده می‌شوند.

۳-۳- ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان (SVM)، یک الگوریتم نظارت‌شده یادگیری ماشین است که برای مسائل طبقه‌بندی و رگرسیون قابل استفاده است. با این حال از آن بیشتر در مسائل طبقه‌بندی استفاده شده است. در الگوریتم ماشین بردار پشتیبان (SVM)، هر نمونه داده به عنوان یک نقطه در فضای n بعدی روی نمودار پراکنندگی داده‌ها ترسیم می‌شود، n تعداد ویژگی‌هایی است که یک نمونه داده دارد و مقدار هر ویژگی مربوط به داده‌ها، یکی از مؤلفه‌های مختصات نقطه روی نمودار را مشخص می‌کند. الگوریتم ماشین بردار پشتیبان (SVM) سعی می‌کند یک سطح یا فضای جداکننده پیدا کند که داده‌های دو کلاس مختلف را با بیشترین فاصله از یکدیگر جدا کند، به این فاصله حاشیه گفته می‌شود. اگر خط جداکننده حاشیه کمی داشته باشد، احتمال طبقه‌بندی نشدن برخی داده‌ها وجود دارد. معمولاً نمونه‌ها به نحوی در فضای ویژگی پخش شده‌اند که طبقه‌بندی خطی قادر به تفکیک آنها نمی‌باشد. در نتیجه ماشین بردار پشتیبان (SVM) از روشی که به آن کرنل گفته می‌شود، استفاده می‌کند. در این روش در واقع توابعی وجود دارند که فضای ورودی بعد پایین را دریافت کرده و آن را به فضای بعد بالاتر تبدیل می‌کنند. این تبدیل، یک مسئله غیر قابل جداسازی را به مسئله قابل جداسازی مبدل می‌کند. به این توابع، کرنل گفته می‌شود. انواع کرنل‌ها عبارتند از کرنل چند جمله‌ای، کرنل گاوسی، کرنل تابع پایه شعاعی و کرنل سیگموئید. در اینجا از کرنل تابع پایه شعاعی استفاده شده که رابطه‌ی آن به صورت زیر می‌باشد:

$$K(x, \hat{x}) = \exp(-\gamma \|x - \hat{x}\|^2) \quad (1)$$

و برای $\gamma > 0$ ، γ با σ نسبت معکوس دارد. همچنین گاهی اوقات γ با استفاده از پارامتر زیر استفاده می‌شود

$$\gamma = 1/2\sigma^2 \quad (2)$$

که در آن واریانس و شباهت یا نزدیکی X و X' بردار ویژگی، محاسبه شده است.

۳-۴- ساختار روش پیشنهادی

روش پیشنهادی در چند گام اقدام به تشخیص وضعیت و حالت راننده می‌کند که در ادامه به تشریح بخش‌های مختلف آن می‌پردازیم.

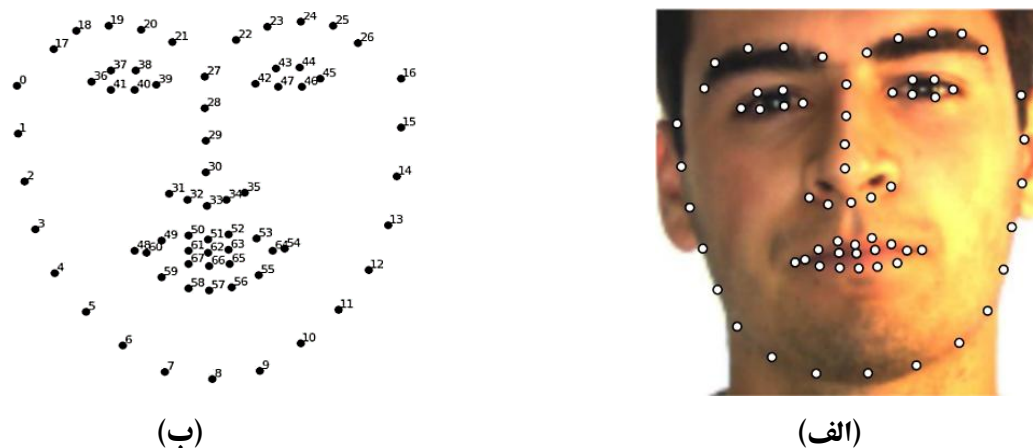
۳-۴-۱- تشخیص چهره در تصاویر رانندگان

در مرحله نخست از روش پیشنهادی، تشخیص چهره راننده به عنوان پیش‌پردازش اصلی انجام می‌شود. برای این کار، از الگوریتم کاظمی و سالیوان [۳۰] که در کتابخانه Dlib پیاده‌سازی شده است، استفاده شده که یکی از دقیق‌ترین و سریع‌ترین الگوریتم‌ها برای تشخیص چهره محسوب می‌شود. این روش مبتنی بر مدل جنگل‌های تصمیم‌گیری تجمعی (Ensemble of Regression Trees) بوده و امکان محلی‌سازی نقاط کلیدی چهره را با دقت بالا فراهم می‌کند. مزیت استفاده از این الگوریتم، سرعت بالای اجرا و عملکرد

مناسب آن حتی در شرایطی مانند تغییرات نور، زوایای مختلف صورت و وجود موانعی مانند عینک یا ماسک جزئی است.

۳-۴-۲- استخراج نقاط کلیدی چهره

پس از تشخیص چهره، الگوریتم کاظمی و سالیوان از مدل ۶۸ نقطه‌ای برای استخراج مختصات نقاط کلیدی چهره استفاده می‌کند. این نقاط کلیدی، مرزهای اصلی اجزای صورت را مشخص کرده و شامل بخش‌هایی مانند چشم‌ها، ابروها، بینی، دهان و فک می‌شود. نگاشت این نقاط روی تصویر چهره راننده، امکان بخش‌بندی دقیق نواحی مختلف صورت را فراهم می‌کند. در شکل ۱، نمونه‌ای از این نقاط کلیدی روی چهره نمایش داده شده است که نشان می‌دهد هر بخش از چهره دارای شاخص‌های مختصات مشخصی است. تجسم ۶۸ مختصات (x, y) برجسته چهره را می‌توان در شکل ۱ (ب) مشاهده کرد.



شکل ۱: شناسایی چهره با کتابخانه Dlib. الف) نمونه‌ای از آشکارسازی ۶۸ مختصات برجسته چهره، ب) تجسم ۶۸ مختصات برجسته چهره.

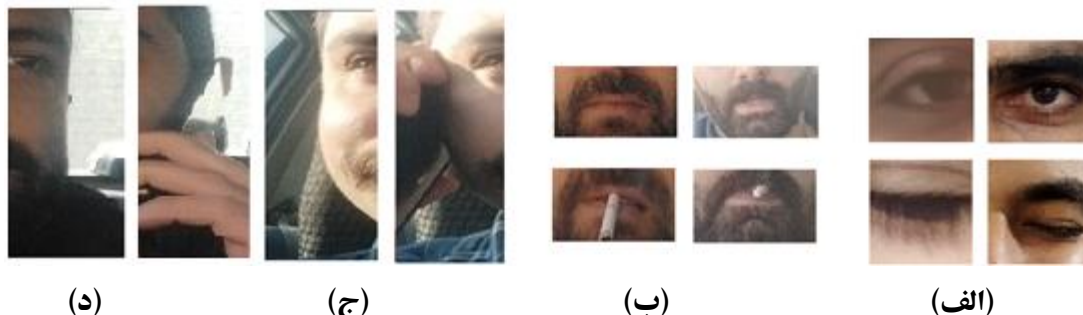
جدول ۱، مقادیر مربوط به بخش‌بندی صورت بر اساس نقاط کلیدی مدل ۶۸ نقطه‌ای Dlib

بخش صورت	محدوده نقاط کلیدی	افست (پیکسل)	توضیحات
چشم چپ	نقاط ۳۶ تا ۴۱	-	شامل پلک بالا و پایین
چشم راست	نقاط ۴۲ تا ۴۷	-	شامل پلک بالا و پایین
ابروی چپ	نقاط ۱۷ تا ۲۱	-	ناحیه ابروی چپ
ابروی راست	نقاط ۲۲ تا ۲۶	-	ناحیه ابروی راست
بینی	نقاط ۲۷ تا ۳۵	-	شامل پل بینی و سوراخ‌های بینی
دهان	نقاط ۴۸ تا ۶۷	-	شامل لب بالا و پایین
گوش چپ	نقاط ۱۷ و ۸	۷۵ پیکسل	تخمین لبه گوش چپ
گوش راست	نقاط ۲۵ و ۸	۷۵ پیکسل	تخمین لبه گوش راست
فک	نقاط ۰ تا ۱۶	-	شامل خط فک و چانه

۳-۴-۳- بخش‌بندی نواحی مختلف صورت

به منظور بررسی وضعیت چهره راننده و تحلیل ویژگی‌های مرتبط با خستگی و حواس‌پرتی، بخش‌بندی دقیق نواحی مختلف چهره انجام می‌شود. برای این منظور، نقاط کلیدی استخراج‌شده از مرحله قبل مورد استفاده قرار گرفته و نواحی به صورت مجزا مطابق جدول ۱ تعیین می‌شوند.

جدول ۱ مشخص می کند که هر بخش صورت چگونه توسط نقاط کلیدی مدل مشخص شده و نواحی اضافی (مانند گوش‌ها) با مقدار آفست ۷۵ پیکسل استخراج شده‌اند. بخش‌بندی به ما این امکان را می‌دهد که هر یک از نواحی چهره را به‌طور جداگانه تحلیل کرده و ویژگی‌های خاص هر بخش را استخراج کنیم. شکل ۲ نشان‌دهنده نحوه جداسازی این نواحی بر اساس نقاط کلیدی و آفست مشخص شده است.



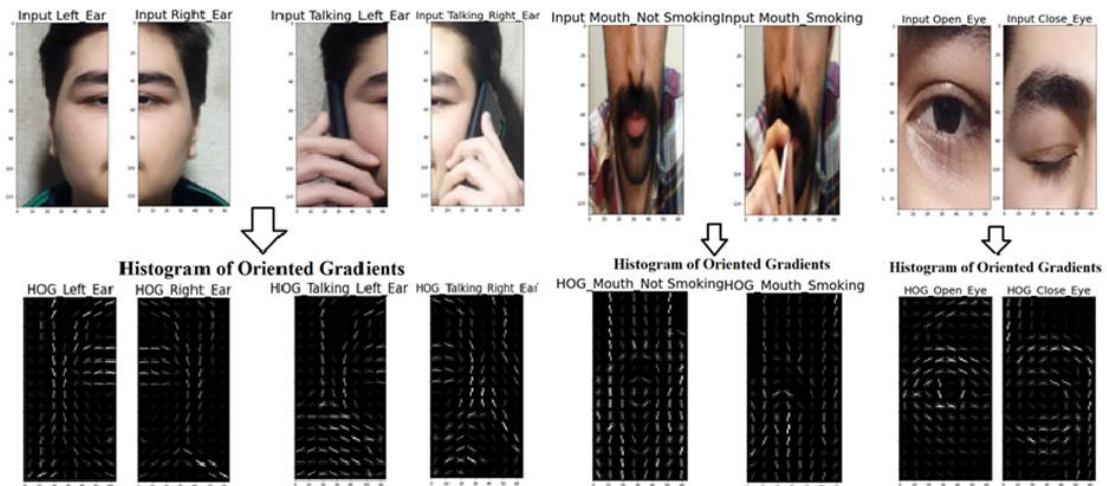
شکل ۲: نمونه جداسازی نواحی مختلف صورت. الف) جداسازی چشم راست و چپ، ب) جداسازی دهان، ج) جداسازی گوش چپ، د) جداسازی گوش راست.

۳-۴-۴- استخراج ویژگی‌های تصویری

پس از جداسازی نواحی مختلف چهره، لازم است ویژگی‌های تصویری استخراج شوند تا امکان تحلیل و طبقه‌بندی رفتار راننده فراهم گردد. در این مرحله، دو روش هیستوگرام گرادیان جهت‌دار (HOG) و الگوی باینری محلی (LBP) برای استخراج ویژگی‌ها مورد استفاده قرار می‌گیرند. قبل از استخراج ویژگی توسط HOG، به عنوان پیش پردازش عکس‌ها به اندازه 128×256 تبدیل می‌شوند. در الگوریتم HOG، تصویر ابتدا به بلوک‌های کوچک تقسیم شده و برای هر بلوک، گرادیان‌های مربوط به تغییرات شدت پیکسل‌ها در جهت‌های مختلف محاسبه و در یک هیستوگرام ذخیره می‌شود. این ویژگی‌ها به‌طور خاص برای شناسایی تغییرات در ساختار چهره، مانند افتادگی پلک‌ها یا تغییر در وضعیت دهان، مفید هستند. برای تشخیص خستگی و حواس‌پرتی رانندگان، ابتدا برای تمامی کلاس‌ها با استفاده از هیستوگرام استخراج ویژگی صورت می‌گیرد. برای هر تصویر آموزشی و تست، ویژگی‌های بدست آمده از روش HOG به صورت یک ماتریس $1 \times n$ می‌باشد که بسته به نوع تصویر ویژگی‌های متفاوتی بدست می‌آید. در شکل ۳ ویژگی‌های استخراج شده از تصاویر و جهت لبه‌ها نشان داده شده است.

ویژگی‌های استخراج شده توسط الگوریتم در چهار حالت برای آموزش به SVM داده می‌شوند. تصاویر مربوطه برچسب گذاری شده و SVM توسط تصاویر آموزش ساخته می‌شود و در ادامه طبقه‌بندی تصاویر آزمایشی توسط SVM صورت می‌پذیرد. در الگوریتم HOG با در نظر گرفتن ۹ جهت گیری^۱ و سلول 8×8 سعی گردیده عوامل حواس‌پرتی و خستگی در رانندگان تشخیص داده شود. شکل ۶ نمونه‌هایی از تشخیص درست و غلط در روش پیشنهادی را نشان می‌دهد. در شکل ۴ نمونه تشخیص‌های صحیح و نادرست از حالت‌های راننده نشان داده شده است.

¹ orientations



(د)

(ج)

(ب)

(الف)

شکل ۳: استخراج ویژگی هیستوگرام گرادیان جهت دار (HOG). الف) استخراج ویژگی چشم راست و چپ، ب) استخراج ویژگی دهان، ج) استخراج ویژگی گوش چپ، د) استخراج ویژگی گوش راست



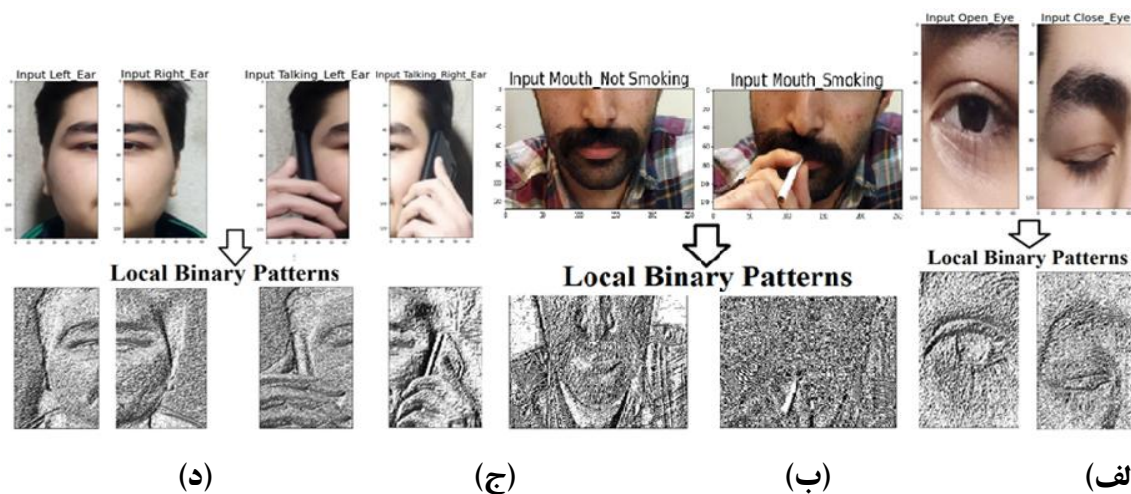
(الف)



(ب)

شکل ۴: نتایج حاصل از استخراج ویژگی و طبقه‌بندی SVM برای تشخیص خستگی و حواس پرتی رانندگان، الف) نمونه‌هایی که درست تشخیص داده شده، ب) نمونه‌هایی که اشتباه تشخیص داده شده.

از روش الگوی باینری محلی LBP، که یکی دیگر از الگوریتم‌های متداول برای تحلیل بافت تصاویر است، در تشخیص ناهنجاری‌ها و تغییرات ظاهری چهره راننده استفاده شده است. در این روش، هر پیکسل از تصویر بر اساس مقدار شدت روشنایی پیکسل‌های همسایه‌اش کدگذاری شده و یک مقدار باینری برای آن محاسبه می‌شود. ویژگی‌های استخراج شده توسط LBP به‌ویژه برای تحلیل تغییرات ظاهری ناشی از خستگی و خواب‌آلودگی، مانند چین و چروک‌های پلک و افتادگی گوشه‌های دهان، بسیار موثر هستند. در شکل ۵ عملکرد الگوریتم بر تمامی کلاس‌ها نشان داده شده است. همانطور که در شکل ۵ نشان داده شده است، بخش‌های چهره شناسایی شده توسط Dlib از مجموعه آموزش استخراج گردیده است. در ادامه تصاویر مربوطه برچسب گذاری شده و از آن‌ها برای ساخت و آموزش SVM روش پیشنهادی استفاده نمودیم. مطابق مراحل الگوریتم HOG، به عنوان پیش پردازش عکس‌ها به اندازه 128×256 تبدیل شده و وارد الگوریتم با چهار کلاس برای آموزش و طبقه بندی می‌شود. در الگوریتم $LBP = 64$ P پیکسل بر روی دایره‌ای به شعاع $R = 4$ حول پیکسل در نظر گرفته شده است. نتایج حاصل از نمونه‌های داده‌های آزمایش در شکل ۶ برای تشخیص صحیح و نادرست حواس‌پرتی و خستگی در رانندگان نشان داده شده است.



شکل ۵: ویژگی‌های استخراج شده توسط الگوریتم الگوی باینری محلی (LBP). الف) استخراج ویژگی چشم راست و چپ، ب) استخراج ویژگی دهان، ج) استخراج ویژگی گوش چپ، د) استخراج ویژگی گوش چپ

۳-۴-۵- طبقه‌بندی رفتار راننده با ماشین بردار پشتیبان (SVM)

پس از استخراج ویژگی‌های مربوط به هر ناحیه از چهره، این ویژگی‌ها به الگوریتم ماشین بردار پشتیبان (SVM) برای طبقه‌بندی داده می‌شوند. SVM یکی از الگوریتم‌های یادگیری نظارت شده است که در دسته‌بندی داده‌ها به دسته‌های مجزا بسیار کارآمد عمل می‌کند. در اینجا، مدل SVM با استفاده از ویژگی‌های استخراج شده از روش‌های HOG و LBP، وضعیت‌های مختلف راننده را تشخیص داده و برچسب گذاری می‌کند.

۳-۴-۶- تشخیص خستگی و حواس‌پرتی راننده

با استفاده از ویژگی‌های استخراج شده و مدل SVM، وضعیت‌های مختلف راننده شناسایی شده و مشخص می‌شود که آیا راننده دچار خستگی یا حواس‌پرتی شده است یا خیر. در این بخش، کلاس‌های مختلف رفتاری راننده شامل چشم باز/ بسته، سیگار کشیدن، نگاه کردن به چپ یا راست، و صحبت کردن با تلفن همراه مورد بررسی قرار می‌گیرند. نتایج نشان می‌دهد که روش پیشنهادی قادر است با دقت بالایی وضعیت‌های مختلف را شناسایی کرده و هشدارهای لازم را ارائه دهد. در بخش نتایج عملکرد روش پیشنهادی به طور کامل بررسی و مورد بحث قرار می‌گیرد.



(الف)



(ب)

شکل ۶: نتایج LBP برای تشخیص خستگی و حواس پرتی رانندگان، (الف) نمونه‌هایی که درست تشخیص داده شده، (ب) نمونه‌هایی که اشتباه تشخیص داده شده.

الگوریتم جداسازی نواحی مختلف صورت

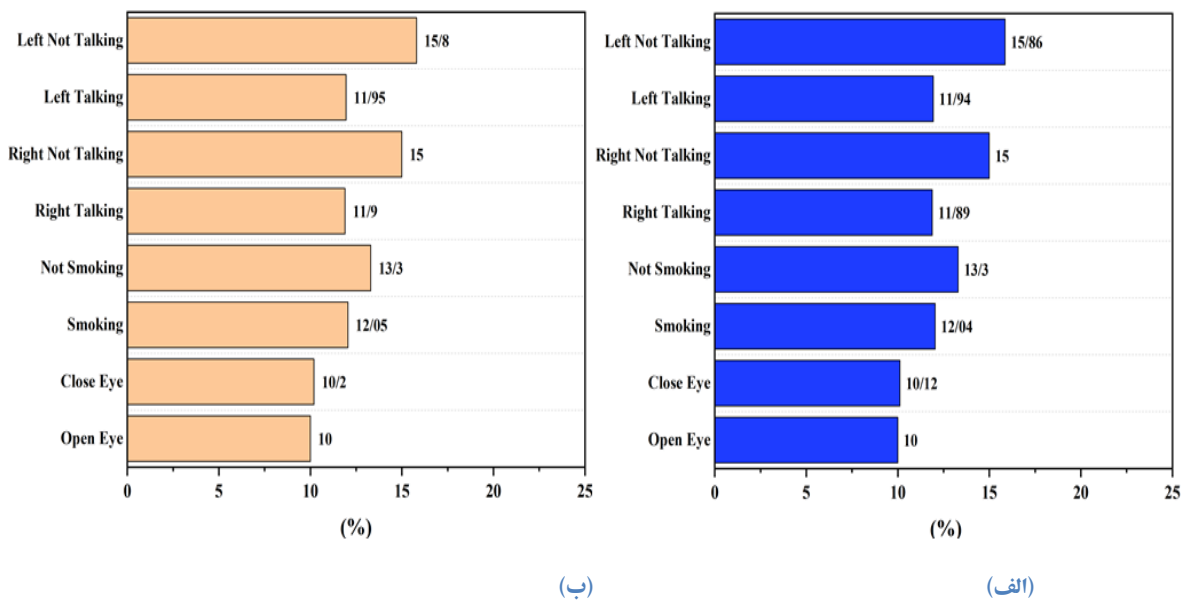
ورودی: تصویر چهره راننده

خروجی: نواحی مختلف چهره (چشم‌ها، بینی، دهان و ...)

۱. بارگذاری مدل تشخیص چهره
۲. دریافت تصویر ورودی و تبدیل آن به سطح خاکستری
۳. اجرای الگوریتم تشخیص چهره روی تصویر
۴. برای هر چهره تشخیص داده شده انجام دهید:
 - ۴.۱. استخراج ۶۸ نقطه کلیدی چهره با استفاده از مدل محلی سازی
 - ۴.۲. استخراج نواحی اصلی چهره:
 - استخراج نقاط مربوط به چشم چپ
 - استخراج نقاط مربوط به چشم راست
 - استخراج نقاط مربوط به بینی
 - استخراج نقاط مربوط به دهان
 - ۴.۳. اعمال آفست‌های مشخص برای جداسازی نواحی چهره
 - ۴.۴. ذخیره یا نمایش نواحی جداسازی شده
۵. پایان

۳-۵- مجموعه داده‌ها

برای ایجاد یک سیستم که خستگی و حواس پرتی رانندگان را تشخیص دهد، در گام نخست نیازمند داده‌هایی از راننده هستیم که منجر به حواس پرتی و خواب‌آلودگی او می‌شود. اکثر پژوهش‌ها از داده‌های منحصر به خودشان و همچنین داده‌هایی که متناسب با زاویه دید ما نبودند استفاده نموده‌اند و اکثر کلاس‌های مورد نظر آن‌ها با کلاس‌های بررسی شده‌ی ما متفاوت بود. در نهایت، مجموعه داده‌ای شامل خواب‌آلودگی، استعمال دخانیات و صحبت با تلفن همراه که با استفاده از تکنیک‌های داده‌افزایی، برهم زدن مقیاس، زاویه چرخش و موقعیت شکل گوش ارتقا پیدا کردند، فراهم گردیده است. تعداد ۲۵۰۶۸ عکس جمع‌آوری گردیده تا به کمک آنها سیستم و طبقه‌بندی بر روی این داده‌ها آموزش داده شود. این مجموعه به دو بخش آموزش و آزمایش تقسیم شده است، بخش آموزش شامل ۲۰۰۵۴ تصویر و بخش آزمایش شامل ۵۰۱۴ تصویر می‌باشد. اطلاعات مربوط به تعداد عکس‌های بخش آموزش و آزمایش به ترتیب برای هر وضعیت در شکل ۷ نشان داده شده است.



شکل ۷: نمودار توزیع داده‌های آموزش و آزمایش (الف) توزیع داده‌های آموزشی بر اساس وضعیت، (ب) توزیع داده‌های آزمایش بر اساس وضعیت.

در شکل ۷ - الف) میزان درصد هر کدام از وضعیت‌های موجود برای بخش آموزش نمایش داده شده است. همانطور که مشاهده می‌شود بیشتر داده‌ها مربوط به صحبت با تلفن همراه و کمترین داده‌ها متعلق به چشم باز و بسته می‌باشد. از آن‌جا که بخش صحبت با تلفن همراه به دو بخش گوش راست و گوش چپ تقسیم می‌شود، لذا توزیع داده‌های آزمایش در شکل مربوط بیشتر است. همچنین در شکل ۷ - ب) میزان درصد هر کدام از وضعیت‌های موجود برای بخش آموزش نمایش داده شده است. در این شکل همانند بخش قبلی، بیشتر داده‌ها مربوط به صحبت با تلفن همراه و کمترین داده‌ها متعلق به چشم باز و بسته می‌باشد. در شکل ۸ حالت‌های مختلف حواس پرتی نشان داده شده است.

حالت‌های مختلف راننده	نمونه تصاویر حالت‌های راننده
حالت عادی	
حالت خواب آلودگی	
حالت استعمال دخانیات	
حالت مکالمه با گوشی در دست چپ	
حالت مکالمه با گوشی در دست راست	

شکل ۸: حالت‌های مختلف حواس پرتی شامل استعمال دخانیات، مکالمه با تلفن همراه و خواب‌آلودگی راننده

۴- نتایج

برای بررسی کارایی روش پیشنهادی، نتایج دقت طبقه‌بندی برای دو روش استخراج ویژگی HOG + SVM و LBP + SVM مقایسه شده‌اند. جدول ۱ نشان‌دهنده دقت به دست آمده در تشخیص وضعیت‌های مختلف است. معیار دقت صحت کلاس‌ها در واقع از تقسیم تعداد کل فریم‌های صحیح مثبت و صحیح منفی بر کل فریم‌ها محاسبه شده است. سپس، از دقت تمامی چهار وضعیت میانگین گرفته شده تا دقت کل برای هر الگوریتم بدست آید. معیار دقت در معادله ۲ نشان داده شده است، که در آن TP تعداد نمونه‌های درست مثبت، TN تعداد نمونه‌های درست منفی، FP تعداد نمونه‌های اشتباه مثبت و FN تعداد نمونه‌های اشتباه منفی می‌باشد.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100 \quad (1)$$

جدول ۲: مقایسه دقت داده‌های آموزش و آزمایش در تشخیص حالت‌ها.

حالت‌ها روش‌ها	چشم باز - بسته	سیگار	گوشی - راست	گوشی - چپ	میانگین (دقت)
LBP + SVM	٪۶۱	٪۶۰	٪۶۲	٪۶۳	۶۱٪/۵
HOG + SVM	٪۸۵	٪۹۰	٪۹۰	٪۸۹	۸۸٪/۵

بر این اساس نتایج حاصل از استخراج ویژگی با استفاده از الگوریتم هیستوگرام گرادیان جهت‌دار (HOG) و الگوریتم الگوی باینری محلی (LBP) با طبقه‌بندی ماشین بردار پشتیبان (SVM) در جدول ۲ بیانگر این حقیقت است که استخراج ویژگی با استفاده از الگوریتم HOG نسبت به استخراج ویژگی الگوریتم LBP در وضعیت چشم باز یا بسته، استعمال دخانیات و صحبت با تلفن همراه به نتایج قابل قبولی دست یافته است. میانگین دقت ٪۸۸.۵ در روش استخراج ویژگی HOG با طبقه‌بندی SVM بهترین عملکرد را به دلیل توانایی بیشتر در استخراج ویژگی‌های ساختاری چهره، برای تشخیص خستگی و حواس‌پرتی در رانندگان را داشته است. روش HOG + SVM به طور کلی عملکرد بسیار بهتری نسبت به LBP + SVM دارد. دقت روش LBP + SVM در محدوده ٪۶۰ تا ٪۶۳ است، که نشان‌دهنده کارایی کمتر این روش در تشخیص حالات مختلف چهره است.

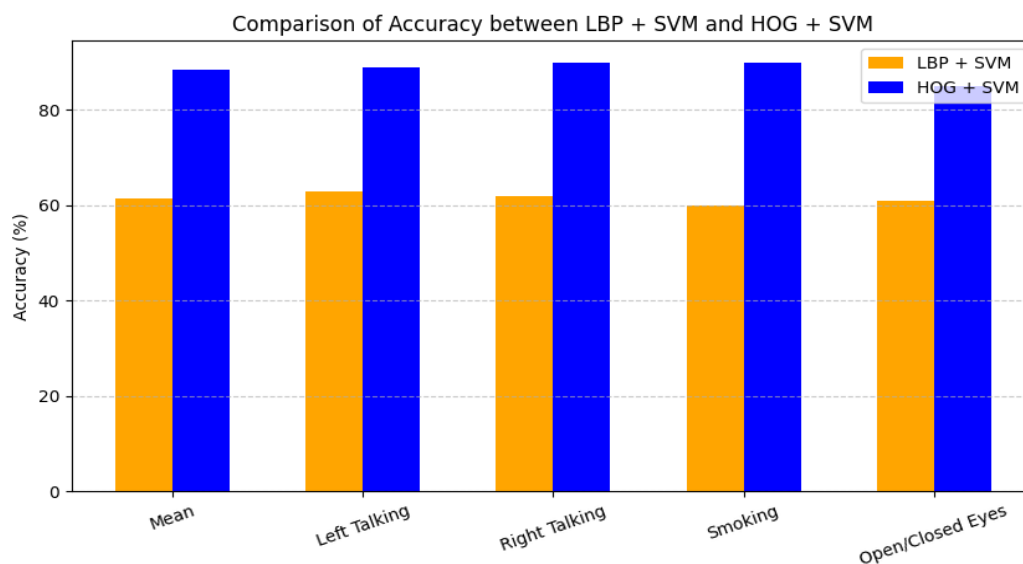
روش HOG + SVM در همه دسته‌ها دقت بالای ٪۸۵ را دارد و برای گوشی راست به دقت ٪۹۰ رسیده است. بهترین عملکرد HOG + SVM برای تشخیص استفاده از گوشی سمت راست (٪۹۰) و سیگار کشیدن (٪۹۰) بوده است. ضعیف‌ترین عملکرد LBP + SVM در تشخیص سیگار کشیدن (٪۶۰) مشاهده می‌شود. ما همچنین جدول ۳ (ارزیابی) را برای Precision (تعداد پیش‌بینی‌های درست از یک کلاس تقسیم بر کل پیش‌بینی‌های انجام شده برای آن کلاس)، Recall (تعداد پیش‌بینی‌های درست از یک کلاس تقسیم بر کل داده‌های واقعی آن کلاس)، F1-Score (میانگین هارمونیک Precision و Recall)، محاسبه نمودیم. بر اساس اطلاعات این جدول، روش HOG + SVM به طور قابل توجهی عملکرد بهتری نسبت به LBP + SVM دارد، به ویژه در تشخیص سیگار و گوشی راست که Precision و Recall بالاتری دارند. مدل LBP + SVM دقت متوسطی

دارد، اما نسبت به HOG + SVM کارایی پایین تری دارد. در هر دو مدل، تشخیص چشم باز و بسته کمترین مقدار دقت را دارد، که نشان دهنده دشواری در شناسایی این ویژگی است.

جدول ۳: مقادیر ارزیابی مدل‌ها در تشخیص حالت‌ها

میانگین	چشم باز - بسته	سیگار	گوشی راست	گوشی چپ	معیارهای ارزیابی (روش)
۰.۶۱	۰.۶۱	۰.۶۰	۰.۶۲	۰.۶۱	Precision (LBP + SVM)
۰.۶۱۵	۰.۶۱	۰.۶۰	۰.۶۲	۰.۶۳	Recall (LBP + SVM)
۰.۶۱۲	۰.۶۱	۰.۶۰	۰.۶۲	۰.۶۲	F1-Score (LBP + SVM)
۰.۸۸۵	۰.۸۵	۰.۹۰	۰.۹۰	۰.۸۹	Precision (HOG + SVM)
۰.۸۸۲	۰.۸۵	۰.۹۰	۰.۹۰	۰.۸۸	Recall (HOG + SVM)
۰.۸۸۳	۰.۸۵	۰.۹۰	۰.۹۰	۰.۸۸۵	F1-Score (HOG + SVM)

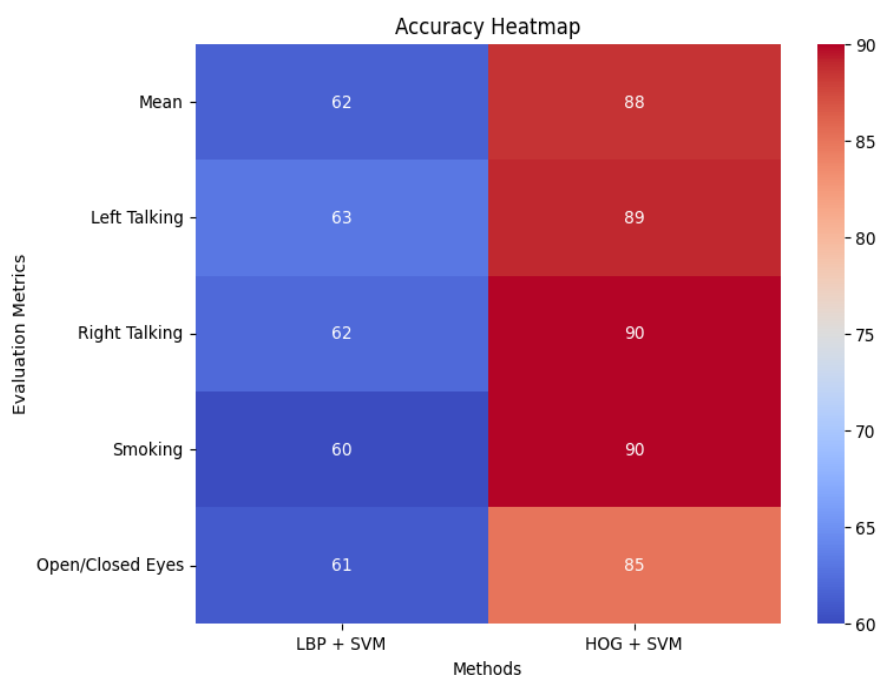
نمودار ارائه شده در شکل ۹، مقایسه‌ای بین دقت دو روش استخراج ویژگی یعنی LBP + SVM و HOG + SVM را در تشخیص حالات مختلف راننده نمایش می‌دهد. محور افقی شامل پنج وضعیت مختلف شامل، میانگین دقت، صحبت کردن با گوشی در سمت چپ، صحبت کردن با گوشی در سمت راست، سیگار کشیدن، و وضعیت باز یا بسته بودن چشم‌ها است. محور عمودی نیز مقدار دقت را به صورت درصد نمایش می‌دهد.



شکل ۹، مقایسه بین دقت دو روش استخراج ویژگی HOG + SVM و LBP + SVM

نتایج نشان می‌دهند که روش HOG + SVM (میل‌های آبی) در تمامی حالات عملکرد بهتری نسبت به LBP + SVM (میل‌های نارنجی) دارد. به طور خاص، مقدار دقت در روش HOG + SVM برای تمامی حالات نزدیک به ۹۰٪ است، در حالی که روش LBP + SVM دقتی در حدود ۶۰ تا ۶۵ درصد دارد. این اختلاف نشان می‌دهد که ویژگی‌های استخراج شده از طریق HOG (هیستوگرام گرادیان‌های جهت‌دار) حاوی اطلاعات بیشتری برای تشخیص الگوهای صورت و حالات راننده هستند و طبقه‌بند SVM

توانایی بالاتری در دسته‌بندی این ویژگی‌ها دارد. به‌ویژه در وضعیت‌های صحبت کردن با گوشی (چپ و راست) و سیگار کشیدن، روش HOG + SVM دقت بالایی را ارائه کرده است، که این نشان‌دهنده قابلیت بالای این روش در تمایز دقیق ویژگی‌های مرتبط با دست و دهان در هنگام استفاده از گوشی یا سیگار است. در مقابل، روش LBP + SVM احتمالاً به دلیل وابستگی به الگوهای بافتی محلی، اطلاعات مهم را در این حالات از دست داده و عملکرد ضعیف‌تری دارد. همچنین، در بخش میانگین دقت، روش HOG + SVM با اختلاف حدود ۲۵٪ نسبت به LBP + SVM برتری دارد، که تأییدی بر کارایی بالای آن در تشخیص تمامی حالات راننده است. همچنین شکل ۱۰ نقشه حرارتی (Heatmap) ارائه شده، دقت دو روش را در پنج معیار ارزیابی مختلف نمایش می‌دهد. محور عمودی شامل میانگین دقت، صحبت کردن با گوشی در سمت چپ، صحبت کردن با گوشی در سمت راست، سیگار کشیدن، و باز یا بسته بودن چشم‌ها است، در حالی که محور افقی نشان‌دهنده دو روش مورد مقایسه یعنی LBP + SVM و HOG + SVM می‌باشد.



شکل ۱۰، نقشه حرارتی مقایسه دقت دو روش در پنج معیار ارزیابی مختلف

رنگ‌بندی نقشه حرارتی به‌وضوح نشان می‌دهد که روش HOG + SVM (ستون دوم) در تمامی معیارها دقت بالاتری نسبت به LBP + SVM دارد. مقدار دقت برای LBP + SVM در محدوده ۶۰٪ تا ۶۳٪ قرار دارد، که نسبتاً پایین است. در مقابل، مقدار دقت برای HOG + SVM بین ۸۵٪ تا ۹۰٪ است که نشان از برتری قابل توجه این روش دارد. بیشترین اختلاف دقت بین دو روش در تمامی معیارها تقریباً ۲۵٪ تا ۳۰٪ است، که تأیید می‌کند روش HOG + SVM ویژگی‌های غنی‌تر و متمایزتری را برای تشخیص حالات مختلف راننده استخراج می‌کند. بهترین عملکرد HOG + SVM در معیار صحبت کردن با گوشی در سمت راست و سیگار کشیدن مشاهده می‌شود که دقتی معادل ۹۰٪ دارد. این نتیجه نشان می‌دهد که روش HOG قادر است ویژگی‌های مهمی از حرکات دست و دهان را که در این دو حالت تأثیرگذارند، بهتر شناسایی کند. کمترین دقت HOG + SVM در معیار باز یا بسته بودن چشم‌ها دیده می‌شود (۸۵٪). این امر ممکن است به دلیل چالش‌های مربوط به تغییرات نور، زاویه دید، و ویژگی‌های چهره باشد.

در مقابل، روش LBP + SVM در تمامی معیارها عملکرد ضعیف‌تری دارد و کمترین دقت آن در معیار سیگار کشیدن با مقدار ۶۰٪ ثبت شده است، که احتمالاً به دلیل ناکارآمدی LBP در استخراج الگوهای پیچیده از ویژگی‌های چهره و اشیای اطراف (مانند

سیگار) است. در جدول ۴ مقایسه‌ای میان روش‌های آزمایش شده، با سایر روش‌های تشخیص خستگی و حواس‌پرتی رانندگان انجام شده است، که در ادامه به تحلیل مقادیر جدول می‌پردازیم.

جدول ۴: مقایسه روش LBP + SVM و HOG + SVM با روش‌های اخیر در تشخیص خستگی و حواس‌پرتی رانندگان.

روش	دقت (%)	سرعت به پردازش (ms)	نیاز داده زیاد	مناسب برای سیستم‌های تعبیه شده	مقاومت در برابر تغییرات محیطی
LBP + SVM	۶۱.۵	۱۲	کم	بله	ضعیف
HOG + SVM	۸۸.۵	۱۸	متوسط	بله	متوسط
Gabor Filters + SVM [31]	۸۴.۳	۲۲	متوسط	بله	متوسط
ResNet-50 [32]	۹۵.۶	۵۵	زیاد	خیر	قوی
CNN + LSTM [33]	۹۶.۸	۶۰	خیلی زیاد	خیر	خیلی قوی
ViT (Transformers) [34]	۹۷.۵	۸۵	بسیار زیاد	خیر	خیلی قوی
MobileNetV2 [35]	۹۲.۱	۳۰	متوسط	بله	قوی
EfficientNet-B3 [36]	۹۴.۷	۴۰	زیاد	خیر	قوی
YOLOv5 [37]	۹۳.۸	۲۵	متوسط	بله	قوی

در این مقایسه، روش پیشنهادی HOG + SVM همچنان به عنوان یک گزینه سریع و کم‌هزینه با دقت ۸۸.۵٪ شناخته می‌شود که عملکرد بهتری نسبت به LBP + SVM (۶۱.۵٪) و Gabor Filters + SVM (۸۴.۳٪) دارد. این روش در پردازش بلادرنگ و سیستم‌های تعبیه شده کارایی مناسبی دارد، اما نسبت به روش‌های یادگیری عمیق از دقت کمتری برخوردار است. همچنین در برابر تغییرات شدید محیطی مانند نورپردازی متغیر و حرکات تصادفی راننده، عملکردی پایین‌تر از مدل‌های عمیق مانند CNN + LSTM و ViT دارد. در میان مدل‌های یادگیری عمیق، ResNet-50 (۹۵.۶٪)، CNN + LSTM (۹۶.۸٪) و ViT (۹۷.۵٪) دقت بالایی ارائه می‌دهند، اما به دلیل نیاز پردازشی زیاد، مناسب سیستم‌های کم‌قدرت نیستند. این امر باعث می‌شود که روش‌های مبتنی بر شبکه‌های عصبی کمتر مناسب برای پردازش بلادرنگ در خودروهای تجاری باشند، مگر اینکه از سخت‌افزارهای پیشرفته مانند GPU یا TPU استفاده شود. EfficientNet-B3 با دقت (۹۴.۷٪) و MobileNetV2 با (۹۲.۱٪)، از نظر پردازشی بهینه‌تر هستند و گزینه‌هایی مناسب برای دستگاه‌های قابل حمل و خودروهای هوشمند محسوب می‌شوند. YOLOv5 نیز به دلیل سرعت پردازش بالا (25ms) و دقت (۹۳.۸٪)، برای کاربردهای تشخیص چهره و حواس‌پرتی رانندگان در ویدئوهای بلادرنگ بسیار مناسب است. این مدل، برخلاف روش‌های دیگر، در رابطه با تشخیص چند کلاسه هم‌زمان (چشم بسته، صحبت با تلفن، سیگار کشیدن و غیره) برتری دارد.

روش پیشنهادی، اگرچه از ویژگی‌های دستی استخراج شده مانند HOG + SVM بهره می‌برد، اما در مقایسه با روش‌های مبتنی بر شبکه‌های عمیق محدودیت‌هایی دارد. شبکه‌های عمیق توانایی یادگیری ویژگی‌های انتزاعی و پیچیده را دارند و با استفاده از تکنیک‌هایی مانند داده‌افزایی و یادگیری تطبیقی، عملکرد بهتری در شرایط متغیر نوری، زاویه‌های مختلف و حرکات تصادفی راننده

نشان می دهند. همچنین، مدل های ترکیبی با تنظیمات پیشرفته ای مانند بهینه سازی چندهدفی و یادگیری تطبیقی دقت تشخیص را افزایش می دهند. برای مثال، مدل های CNN-LSTM قادر به بررسی تغییرات زمانی رفتار راننده هستند، درحالی که روش HOG + SVM تنها ویژگی های استاتیک چهره را تحلیل می کند. علاوه بر این، ترکیب روش های مختلف مانند PCA + HOG + SVM یا استفاده از چند توصیف گر ویژگی می تواند عملکرد بهتری ارائه دهد. یکی از چالش های دیگر روش پیشنهادی، نیاز به تنظیم دقیق پارامترها (مانند اندازه پنجره و تعداد باین های هیستوگرام) است، درحالی که شبکه های عمیق می توانند این ویژگی ها را به صورت خودکار بهینه سازی کنند. در مجموع، روش HOG + SVM یک توازن بهینه بین دقت، سرعت پردازش و هزینه محاسباتی ارائه می دهد. این روش می تواند به عنوان جایگزینی کم هزینه و سریع برای سیستم های هشداردهنده خستگی راننده در خودروهای هوشمند به کار رود، به ویژه در مواردی که منابع سخت افزاری محدود هستند. برای تحقیقات آینده، پیشنهاد می شود از مدل های هیبریدی که ترکیبی از شبکه های عمیق و روش های کلاسیک استخراج ویژگی هستند، استفاده کنیم، تا ضمن حفظ قابلیت تفسیرپذیری روش های کلاسیک، دقت و پایداری در شرایط مختلف افزایش یابد. ترکیب این روش با یادگیری عمیق هیبریدی (مانند HOG + CNN)، یک CNN سبک (مانند MobileNetV2) و یا استفاده از YOLOv5 برای تحلیل ویدیویی می تواند دقت را افزایش داده و نقاط ضعف آن را کاهش دهد.

به طور کلی نمودار نقشه حرارتی و اطلاعات جداول این تحقیق نشان دادند که روش HOG + SVM به طور قابل توجهی از روش LBP + SVM بهتر عمل می کند، به ویژه در تشخیص حالات پیچیده مانند صحبت کردن با گوشی و سیگار کشیدن. این نتایج اهمیت استفاده از ویژگی های گرادینانی در استخراج الگوهای چهره و حرکات را برجسته می کند و پیشنهاد می شود که برای کاربردهای مرتبط با تشخیص حواس پرتی راننده، روش HOG + SVM به عنوان گزینه بهینه در نظر گرفته شود و می تواند تأثیر مستقیمی بر بهبود دقت طبقه بندی داشته باشد.

۵- نتیجه گیری

تشخیص خستگی و حواس پرتی رانندگان یکی از چالش های اساسی در افزایش ایمنی جاده ای است. طبق گزارش سازمان امنیت ترافیک بزرگراه ملی (NHTSA) در آمریکا، حواس پرتی و خواب آلودگی یکی از عوامل تصادفات برون جاده ای می باشد. حواس پرتی باعث کاهش آگاهی و تصمیم گیری رانندگان شده، در نتیجه پاسخ به رویدادهای رانندگی با تاخیر و اختلال در سرعت همراه بوده است. به همین منظور نیاز به سیستم تشخیص خواب آلودگی و عوامل منجر به حواس پرتی احساس می شود که بتواند به صورت بلادرنگ در کوتاهترین زمان ممکن عامل حواس پرتی راننده را تشخیص و هشدار دهد. در این پژوهش، سیستمی مبتنی بر پردازش تصویر توسعه داده شد که شامل مراحل تشخیص چهره، بخش بندی اجزای چهره و طبقه بندی وضعیت راننده است. روش های مختلفی برای این منظور مقایسه شدند که در این میان، HOG + SVM با دقت ۸۸.۵ درصد، به عنوان گزینه ای سریع و کم هزینه عملکرد بهتری نسبت به LBP + SVM با دقت ۶۱.۵ درصد داشت، اما همچنان در برابر تغییرات شدید محیطی، نسبت به روش های مبتنی بر یادگیری عمیق محدودیت هایی دارد. مدل های عمیقی مانند ViT و CNN + LSTM دقت بالاتری ارائه می دهند، اما نیاز پردازشی بالای آنها باعث می شود برای سیستم های تعبیه شده مناسب نباشند. برای تحقیقات آینده، استفاده از مدل های هیبریدی که ترکیبی از ویژگی های کلاسیک و شبکه های عصبی هستند، می تواند به بهبود دقت و پایداری سیستم در روش پیشنهادی کمک کند. پیاده سازی این سیستم در Python و اجرای آن در Google Colab انجام شد.

- [1] Transportation, U.S.D, "National Highway Traffic Safety Administration". 2022; Available from: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813266>
- [2] Overview of Motor Vehicle Crashes in 2020, "National Highway Traffic Safety Administration". 2018; Available from: <https://www.nhtsa.gov/campaign/distracted-driving>
- [3] H.Veeraraghavan, N. Bird, S. Atev, et al., "Classifiers for driver activity monitoring", Transportation Research Part C: Emerging Technologies, Vol. 15, Issue. (1), pp. 51-67, Jan. 2007.
- [4] A. Azman, Q. Meng, E. A. Edirisinghe, et al., "Eye and mouth movements extraction for driver cognitive distraction detection", in 2012 IEEE Business, Engineering & Industrial Applications Colloquium (BEIAC), Apr. 2012.
- [5] Q. Ji, Z. Zhu, P. Lan, Real-time nonintrusive monitoring and prediction of driver fatigue", IEEE Transactions on Vehicular Technology, Vol. 53, Issue. (4), pp. 1052-1068, Jul.2004.
- [6] F. Vicente, Z. Huang, T. Tran, et al., "Driver gaze tracking and eyes off the road detection system", IEEE Transactions on Intelligent Transportation Systems, Vol. 18, Issue. (4), pp. 837-848, Aug. 2015.
- [7] A. Doshi, M. M. Trivedi, "Head and eye gaze dynamics during visual attention shifts in complex environments", Journal of Vision, Vol. 12, Issue. (2), pp. 9-10, Feb. 2012.
- [8] Y. Abouelnaga, H. M. Eraqi, M. N. Moustafa, "Real-time distracted driver posture classification", Conference: 2018 NIPS Workshop on Machine Learning for Intelligent Transportation SystemsAt: Montreal, Canada, Vol. 20, Issue. (5), pp. 1572-1582. Dec. 2018.
- [9] B. Baheti, S. Gajre, S. Talbar, "Detection of Distracted Driver Using Convolutional Neural Network", Conference: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1-5, Jun. 2018.
- [10] M. U. Khan, T. Mahmood, A. Afsar, "Vision Transformers for Driver Distraction Detection", Neural Computing and Applications, pp. 1-15, Aug. 2022.
- [11] S. Jung, S. Jo, C. H. Im, "Brain activity classification for driver drowsiness detection using deep learning with EEG signals", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 26, Issue. (7), pp. 1239-1245, Feb. 2019.
- [12] C. T. Lin, L. W. Ko, I. F. Chung, et al., "Adaptive EEG-based alertness estimation system by using ICA-based fuzzy neural networks", IEEE Transactions on Circuits and Systems I: Regular Papers, Vol. 53, Issue. (11), pp. 2469-2476. Des. 2006.
- [13] Y. Liu, Y. Wang, C. H. Wu, et al., "Multi-modal sensor fusion for detecting driver distraction", IEEE Transactions on Intelligent Vehicles, Vol. 4, Issue. (2), pp. 179-188, Jul. 2019.
- [14] N. George, S. Ghosh, H. Zhang, et al., "A multi-modal approach to driver distraction detection using deep learning", IEEE Transactions on Intelligent Transportation Systems, Vol. 22, Issue. (3), pp. 1478-1490. Dec. 2021.
- [15] S. K. L. Lal, A. Craig, "A critical review of the psychophysiology of driver fatigue", Biological Psychology, Vol. 55 Issue. (3), pp.173-194, Feb. 2001.
- [16] M. Awais, N. Badruddin, M. Drieberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability", Sensors, Vol. 17, Issue. (9), pp. 1991, Aug. 2017.
- [17] B. T. Jap, S. Lal, P. Fischer, et al. , "Using EEG spectral components to assess algorithms for detecting fatigue", Expert Systems with Applications, Vol. 36, Issue. (2), pp. 2352-2359. Mar. 2009.
- [18] H. Singh, N. Papanikolopoulos, "Monitoring driver distraction using facial tracking and head pose estimation", IEEE Transactions on Intelligent Transportation Systems, Vol. 16, Issue. (2), pp. 636-650. Dec. 2015.
- [19] R. A. Naqvi, M. Arsalan, G. Batchuluun, et al., "Deep Learning-Based Gaze Detection System for Automobile Drivers Using a NIR Camera Sensor", Sensors, Vol. 18, Issue. (2), pp. 456, Feb. 2018.
- [20] M. Suleman, S. B. Zulifqar, N. Memon, "Real-time driver drowsiness detection using computer vision

- techniques”, International Electrical Engineering Conference, Apr. 2021.
- [21] A. Mobin, M. M. Rahman, M. A. Rahman, ” Real-time image processing-based drowsiness detection using eye aspect ratio and facial landmark points”, International Conference on Computer and Information Technology (ICCIT), pp.245-250, Jul. 2019.
- [22] M. Abdulmalik, A. Rahman, , A. A. Bakar, ” Facial monitoring system for drowsiness detection using machine learning”, Journal of Intelligent & Fuzzy Systems, Vol. 40, Issue. (1), pp.1-11, Feb. 2021.
- [23] R. Petoucha, W. Zhang, M. Rezaei, ” A computer vision-based approach to detect distracted driving behaviors”, IEEE Transactions on Intelligent Vehicles, Vol. 5, Issue. (4), pp. 512-522, Jul. 2020.
- [24] J. H. Kim, S. H. Bae, ” Smoke Detection Method Using Local Binary Pattern Variance in RGB Contrast Image”, Journal of Korea Multimedia Society, Vol. 18, Issue. (10), pp. 1197-1204, Oct. 2015.
- [25] L. Shaoji, , C. Xiong, ” Deep learning-based driver activity recognition for distraction detection”, Neural Computing and Applications, Vol. 33, Issue. (5), pp. 1749-1762, Oct. 2021.
- [26] T. Ojala, M. Pietikainen, T. Maenpaa, ” Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”, IEEE Transactions on pattern analysis and machine intelligence, Vol. 24, Issue. (7), pp. 971-987, Jul. 2002.
- [27] M. Arif , KB. Khattak, K. Fiaz, et al., ” A Real-Time Driver Drowsiness Detection and Warning System Based on an Eye Blinking Rate”, in Intelligent Technologies and Applications: Second International Conference, INTAP 2019, Bahawalpur, Pakistan, pp. 106-117, May. 2020.
- [28] Adrian Rosebrock. Detect eyes, nose, lips, and jaw with dlib, OpenCV, and Python. 2021; Available from: <https://pyimagesearch.com/2017/04/10/detect-eyes-nose-lips-jaw-dlib-opencv-python/>
- [29] N. Dalal, B. Triggs. ” Histograms of oriented gradients for human detection”, in 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), Jul. 2005.
- [30] V. Kazemi, J. Sullivan. ” One millisecond face alignment with an ensemble of regression trees”, in Proceedings of the IEEE conference on computer vision and pattern recognition, Jun. 2014.
- [31] C. Liu, H. Wechsler, ”Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition”, IEEE Transactions on Image Processing, Vol. 11, Issue. (4), Apr. 2002.
- [32] K. He, X. Zhang, S. Ren, J. Sun, ”Deep Residual Learning for Image Recognition”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Dec. 2016.
- [33] J. Donahue, L. A. Hendricks, S. Guadarrama, et al., ” Long-term recurrent convolutional networks for visual recognition and description”, IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 39, Issue. (4), pp. 677 – 691, Apr. 2017.
- [34] A. Dosovitskiy, L. Beyer, A. Kolesnikov, et al., ” An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale”, International Conference on Learning Representations, Oct. 2020.
- [35] M. Sandler, A. Howard, M. Zhu, et al., ” MobileNetV2: Inverted Residuals and Linear Bottlenecks”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Dec. 2018.
- [36] M.Tan, Q. Le, ” EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”, International Conference on Machine Learning (ICML), Vol. 97, pp. 6105-6114, May. 2019.
- [37] G. Jocher, A. Stoken, A. Chaurasia, et al., "YOLOv5 by Ultralytics." GitHub Repository.



Sahand University
of Technology

DOR:

[20.1001.1.23223146.1404.12.1.3.6](https://doi.org/10.1001.1.23223146.1404.12.1.3.6)

Journal of Nonlinear
Systems in Electrical
Engineering

Vol.12, No.1

Spring and Summer 2025

ISSN: 2322 – 3146

<http://jnsee.sut.ac.ir>

Mobile Robot Localization in Indoor Environments Using a Hybrid Extended Kalman Filter and Rapidly-Exploring Random Tree Algorithm

Mohammadrasoul Haj Ali, Ramazan Havangi*

Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran

Email: havangi@birjand.ac.ir

Abstract

This paper presents an integrated framework for mobile robot navigation in indoor environments, based on the fusion of the Extended Kalman Filter (EKF) and the Rapidly-exploring Random Tree* (RRT) algorithm. The objective of this hybrid approach is to simultaneously enhance path planning accuracy and localization precision under conditions of sensor noise and dynamic, uncertain obstacles. Initially, the RRT algorithm generates an optimal path from the start to the goal point, considering existing obstacles. Subsequently, for high-accuracy path following, the robot's position is estimated in real-time using the EKF. In this method, the prediction step of the filter relies on the robot's kinematic model and odometry data, while the update step is performed using Lidar sensor measurements. The precise localization provided by the EKF enables the robot to follow the proposed RRT* path with minimal error. A key feature of this framework is the closed-loop interaction between localization and path planning; if the estimation error increases or unexpected environmental changes occur, the path planning process is reactivated by the RRT* module. This design increases the system's adaptability and stability in the face of environmental uncertainties. Simulation results demonstrate that this framework significantly improves localization accuracy, achieving a Root Mean Square Error (RMSE) of less than 0.055 and a 73% increase in precision compared to pure odometry. Furthermore, the paths generated by RRT* show considerable improvement in length, smoothness, and safety over those produced by the basic RRT algorithm.

Keywords

Extended Kalman Filter; Indoor Environments; Sampling-based Algorithms; Rapidly-exploring Random Tree; Mobile Robot Localization.

1. Short Introduction

Accurate localization of mobile robots in unknown or dynamic environments remains a fundamental challenge in autonomous navigation. While common methods like odometry suffer from cumulative error, and advanced filters like EKF improve estimation, they often operate independently of path planning. Similarly, path planners like RRT* may generate inefficient routes if the robot's positional uncertainty is high. This paper is motivated by the need to bridge this gap, proposing a novel closed-loop EKF-RRT* framework that synergistically integrates precise, real-time state estimation with reactive and optimal path planning to enhance overall navigation robustness in complex indoor settings.

2. Proposed Work

The proposed system establishes a closed-loop integration between an Extended Kalman Filter (EKF) for localization and a Rapidly-exploring Random Tree (RRT*) algorithm for path planning. The EKF fuses noisy wheel odometry data with Lidar measurements to provide a reliable real-time pose estimate and, crucially, a measure of estimation uncertainty (covariance). The RRT module uses this estimated position to generate an optimal, collision-free path. The core innovation is the feedback link: if the positional uncertainty from the EKF exceeds a threshold—indicating potential localization drift or new obstacles—the system triggers the RRT* to replan a new path from the robot's current estimated location. This ensures the path remains feasible despite uncertainties.

Simulation results in indoor environments validate the framework's superiority. The EKF-based localization achieved a Root Mean Square Error (RMSE) of 0.055, outperforming pure odometry where error grew unbounded. The closed-loop interaction produced paths that were not only collision-free but also 73% more accurate in tracking compared to odometry. The paths were shorter and smoother than those from a standard RRT, as frequent replanning from accurate poses led to more efficient tree growth and obstacle avoidance.

3. Conclusion

This paper successfully developed and validated a hybrid EKF-RRT* framework for robust mobile robot navigation. The closed-loop design synergizes high-precision localization with adaptive path planning, where each module informs and improves the other. Key results demonstrate a 73% increase in localization accuracy over odometry and the generation of safer, more optimal paths. The proposed system represents a significant step toward reliable autonomous navigation for service, industrial, and rescue applications in uncertain indoor environments.



دانشگاه صنعتی شاهرود

DOR:

20.1001.1.23223146.1404.12.1.3.6

نشریه سالانه علمی غیرخطی در مهندسی برق

دوره ۱۲ - شماره ۱

بهار و تابستان ۱۴۰۴

صفحات ۵۷ الی ۸۰

ISSN: 2322-3146

http://jnsee.sut.ac.ir

موقعیت یابی ربات متحرک در محیط های داخلی با استفاده از ترکیب

فیلتر کالمن توسعه یافته و الگوریتم مسیریابی درخت تصادفی

محمد رسول حاج علی، رمضان هاوونگی*

دانشکده مهندسی برق و کامپیوتر - دانشگاه بیرجند - بیرجند - ایران

نویسنده مسوول havangi@birjand.ac.ir

تاریخ پذیرش: ۱۴۰۴/۴/۲۴

تاریخ بازنگری: ۱۴۰۴/۳/۲۳

تاریخ ارسال: ۱۴۰۴/۱/۱۸

چکیده

واژه‌های کلیدی

فیلتر کالمن توسعه یافته،

محیط های داخلی،

الگوریتم های مبتنی بر نمونه برداری،

الگوریتم کاوش سریع درخت

تصادفی،

موقعیت یابی ربات متحرک.

در این مقاله، یک چارچوب یکپارچه برای ناوبری ربات متحرک در محیط های داخلی ارائه شده است که مبتنی بر ترکیب فیلتر کالمن توسعه یافته (EKF) و الگوریتم درخت تصادفی با کاوش سریع (RRT*) می باشد. هدف از این ترکیب، بهبود همزمان دقت مسیریابی و موقعیت یابی در شرایطی است که داده های حسگر دارای نویز هستند و محیط دارای موانع پویا و نامعین می باشد. در گام اول، RRT* مسیر بهینه ای از نقطه شروع تا هدف را با در نظر گرفتن موانع موجود ایجاد می کند. سپس، برای دنبال کردن دقیق مسیر، تخمین موقعیت ربات در طول حرکت با استفاده از EKF انجام می شود. در این روش، مرحله پیش بینی فیلتر مبتنی بر مدل حرکتی ربات و داده های ادومتری است و مرحله به روزرسانی از طریق اندازه گیری های حسگر لیدار صورت می گیرد. موقعیت یابی دقیق به دست آمده از EKF، به ربات امکان می دهد که مسیر پیشنهادی RRT* را با خطای حداقلی دنبال کند. نکته کلیدی در این چارچوب، ارتباط حلقه بسته بین موقعیت یابی و مسیریابی است؛ به طوری که در صورت افزایش خطای تخمین یا تغییرات غیرمنتظره در محیط، فرآیند برنامه ریزی مسیر مجدداً توسط RRT* فعال می شود. این طراحی باعث افزایش تطبیق پذیری و پایداری سیستم در مواجهه با عدم قطعیت های محیطی می گردد. نتایج حاصل از شبیه سازی های متعدد نشان می دهد که این چارچوب، توانسته است دقت موقعیت یابی را به میزان قابل توجهی بهبود دهد. میانگین خطای ریشه میانگین مربعات (RMSE) موقعیت یابی در آزمایش ها کمتر از ۰/۰۵۵ گزارش شده است. همچنین، در مقایسه با الگوریتم های پایه مانند ادومتری صرف، دقت موقعیت یابی به میزان ۷۳٪ افزایش یافته است. مسیرهای تولید شده توسط RRT* نیز از نظر طول، همواری و ایمنی نسبت به مسیرهای ایجاد شده توسط نسخه اولیه RRT* بهبود چشم گیری داشته اند. این نتایج نشان می دهد که ادغام EKF و RRT* نه تنها به بهبود عملکرد هر کدام از این ماژول ها منجر شده، بلکه به طراحی یک سیستم ناوبری هوشمند و مقاوم کمک کرده است که قابلیت استفاده در کاربردهای ربات های خدماتی، صنعتی و امدادی را داراست.

۱- مقدمه

موقعیت یابی دقیق ربات های متحرک، به ویژه در محیط های ناشناخته یا پویا، یکی از چالش های اساسی در حوزه ناوبری خودکار به شمار می رود. ربات های متحرک در دهه های اخیر دستاوردهای قابل توجهی در حوزه هایی مانند تولید صنعتی، علوم پزشکی، کشاورزی، آموزش و تحقیقات فضایی داشته اند. برای تحقق ناوبری ایمن و رسیدن به اهداف کنترلی، ربات باید از دقت بالایی در موقعیت یابی برخوردار باشد. در یک سیستم ناوبری معمول، ربات مراحل مختلفی شامل ادراک، موقعیت یابی، شناخت و کنترل حرکت را طی می کند. در مرحله ادراک، داده های سنسورها به اطلاعات معنادار تبدیل می شوند؛ در مرحله موقعیت یابی، ربات با استفاده از داده های حسگرهای داخلی و خارجی موقعیت خود را در فضای کاری تخمین می زند؛ در مرحله شناخت، مسیر حرکت برنامه ریزی می شود و در نهایت، با کنترل حرکت، خروجی های موتورهای تنظیم می شوند تا ربات به مقصد برسد. در این میان، مرحله موقعیت یابی از اهمیت ویژه ای برخوردار است و در سال های اخیر مورد توجه قابل توجهی از سوی پژوهشگران حوزه رباتیک و سیستم های هوشمند قرار گرفته است [۱]. یکی از روش های متداول در این زمینه، الگوریتم ادومتری است که در ربات های متحرک به کار می رود [۲-۴]. با این حال، خطای تجمعی ناشی از ادومتری با افزایش مسافت طی شده افزایش می یابد و منجر به عدم قطعیت در موقعیت یابی می شود. الگوریتم های فیلتر کالمن توسعه یافته (EKF) به واسطه ساختار بازگشتی خود می توانند با به روزرسانی های پیوسته خطا را کاهش داده و دقت ناوبری را بهبود بخشند [۵].

در سال های اخیر، تعداد قابل توجهی از مطالعات در حوزه قابلیت های ناوبری خودکار و پایداری ربات های متحرک منتشر شده است. برای مثال، در [۶] از رویکرد مبتنی بر فیلتر کالمن توسعه یافته (EKF) در مدلی که از حسگر سونار استفاده می کند برای موقعیت یابی ربات متحرک بهره گرفته شده است. همچنین در [۷]، رویکرد فیلترینگ مبتنی بر EKF به کار گرفته شده است که از نقشه به عنوان منبع کمکی برای موقعیت یابی ربات های متحرک استفاده می کند. در [۸]، مقایسه ای بین فیلتر کالمن خطی (KF) و EKF ارائه شده است که تحلیل شبیه سازی ها نشان می دهد EKF در تخمین موقعیت ربات متحرک به طور قابل توجهی بهتر عمل می کند. در این رابطه، لازم به ذکر است که فیلتر کالمن خطی که در سیستم هایی با دینامیک خطی کاربرد دارد، در سیستم های غیرخطی نیز با اعمال اصلاحاتی قابل استفاده است؛ به این صورت که در هر مرحله از تخمین حالت، با استفاده از سری تیلور، دینامیک سیستم غیرخطی خطی سازی می شود تا بتوان از همان فیلترهای کالمن خطی بهره گرفت [۹]. در این منبع همچنین نشان داده شده است که الگوریتم فیلتر کالمن خطی تا حدی قابلیت تطبیق با شرایط غیرخطی سیستم را دارد. باید توجه داشت که اغلب فیلترهای کالمن به صورت گسسته پیاده سازی می شوند.

در [۱۰]، یک تکنیک ناوبری بصری بدون رانش برای موقعیت یابی ربات متحرک ارائه شده است که در آن فناوری موقعیت یابی باند فوق وسیع (UWB) به عنوان منبع داده کمکی استفاده می شود. همچنین در [۱۱] یک روش موقعیت یابی با استفاده از EKF برای ترکیب داده های حاصل از دو حسگر مختلف، یعنی ادومتری و بینایی کامپیوتری پیاده سازی شده است. نتایج موقعیت یابی در محیط شبیه سازی شده با نقاط شروع شناخته شده و ناشناخته ارائه شده است؛ با این حال، مسئله برنامه ریزی مسیر در این مطالعه مورد بررسی قرار نگرفته است. در [۱۲] تمرکز بر مکان یابی دقیق مبتنی بر شبکه حسگر بی سیم (WSN) برای ربات متحرک است که در

آن با استفاده از شناسایی شرایط عدم دید مستقیم (NLOS) تلاش شده خطاهای بزرگ تر ناشی از روش های سنتی کاهش یابد. به منظور حل این مشکل، توزیع موقعیت محاسبه شده از چند اندازه گیری مختلف برای تخمین مکان ربات متحرک استفاده شده است. در [۱۳] به بررسی موقعیت یابی ربات متحرک بر اساس الگوریتم EKF و یک نقشه مبتنی بر ویژگی های زوایا پرداخته شده است. در این مقاله، زوایای گوشه به عنوان ویژگی های محیط شناسایی و فرآیند استخراج دقیق این ویژگی ها شرح داده شده است. همچنین مدل حرکتی ربات و داده های ادومتری توضیح داده شده و الگوریتم EKF به طور کامل ارائه شده است. در [۱۴]، موقعیت یابی یک ربات رهبر و یک ربات دنبال کننده بر اساس ترکیب داده های حسگرهای مختلف ارائه شده است. در این مطالعه، ربات رهبر مجهز به یک حسگر در بالای خود است که نشانگرهایی با چهار شناسه مختلف بر روی آن نصب شده اند. ربات دنبال کننده دارای پلتفرم چرخش با یک درجه آزادی است که حسگر اینتل روی آن نصب شده است. داده های رمزگذار چرخ و حسگرهای IMU در هر دو ربات با استفاده از EKF ترکیب می شوند تا وضعیت ربات ها به دست آید. کنترل کننده PID حسگر اینتل روی ربات دنبال کننده را می چرخاند تا نشانگرهای ربات رهبر را دنبال کند و موقعیت ربات رهبر پس از شناسایی نشانگرها محاسبه می شود. از طرفی، در زمینه برنامه ریزی مسیر، در [۱۵] الگوریتم RRT* برای جمع آوری واگن برقی چمدان رباتیک در فرودگاه ها به کار رفته است. این مقاله نشان می دهد که الگوریتم های برنامه ریزی حرکت مبتنی بر ریسک موجود به طور موثری خطرات برخورد با موانع استاتیک و پویا را کاهش می دهند. الگوریتم RRT* برخلاف الگوریتم های سنتی که فقط به یک درخت ریشه دار یا درخت های دوتایی تکیه دارند، چندین درخت فرعی مستقل را شامل می شود که هر کدام به صورت جداگانه محیط اطراف خود را بررسی می کنند. در همان زمان، درخت ریشه دار اولیه اطلاعات اکتشافی را از این درخت های فرعی جمع آوری و پیشرفت سریع به سمت وضعیت هدف را تسهیل می کند. ارزیابی های انجام شده، شامل شبیه سازی ها و مطالعات محیطی در دنیای واقعی، نشان می دهد که RRT* از نظر کارایی و استحکام در برنامه ریزی مسیر بهتر از الگوریتم های مبتنی بر ریسک یک طرفه و دو طرفه عمل می کند.

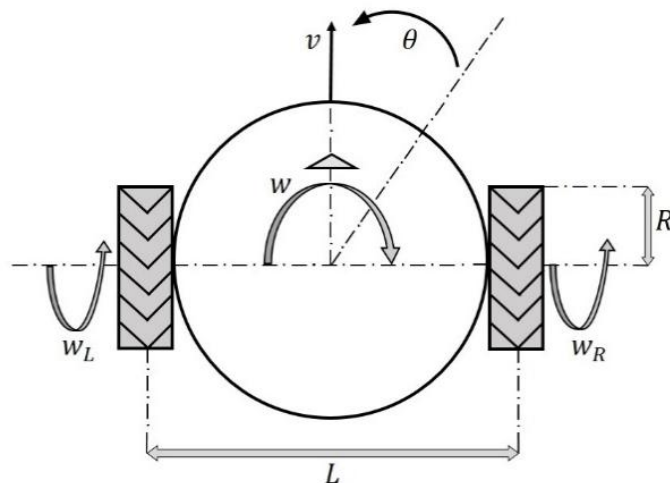
در [۱۶]، روش جدیدی برای برنامه ریزی مسیر ارائه شده است که عدم قطعیت حرکت در محیط های طبیعی را در نظر می گیرد. در این روش، ذرات برای انتشار عدم قطعیت در محیط های پیچیده و با انواع زمین مختلف استفاده می شوند. الگوریتم RRT* بر اساس میزان عدم قطعیت هر گره به منظور جلوگیری از انباشت افزایش عدم قطعیت موقعیت گسترش یافته است. در نتیجه، مسیر تولید شده زمان پیگیری مسیر و برنامه ریزی مجدد را بر اساس اطلاعات موقعیت یابی نادرست کاهش می دهد. نتایج نشان می دهد که روش پیشنهادی عدم قطعیت موقعیت را کاهش داده و همزمان احتمال اجتناب از برخورد و رسیدن به منطقه هدف را در مقایسه با روش های مرسوم حفظ می کند.

در [۱۷] به محدودیت های الگوریتم RRT سنتی در برنامه ریزی مسیر بازوی رباتیک پرداخته شده است، از جمله عملکرد تصادفی، سرعت برنامه ریزی آهسته، مسیرهای غیرهموار و گوشه های بیش از حد. در این مطالعه، الگوریتم RRT بهبود یافته ای با استراتژی سوگیری هدف و رویکرد میدان پتانسیل مصنوعی پیشنهاد شده است. از نظر کاربردهای ناوبری هوایی، در [۱۸] فیلتر کالمن مقاوم با محدودیت حرکتی برای مکان یابی ربات در محیط هوایی که با انسداد سیگنال رادیویی و ویژگی های بصری مشخص می شود، توسعه یافته است. این مقاله عملکرد بهینه EKF را در کاربردهای ناوبری هوایی نشان می دهد. علاوه بر این، در [۱۹] الگوریتم ES-EKF برای موقعیت یابی بلادرننگ در محیط های فاقد GNSS پیشنهاد شده است که سیستم پیشنهادی را در برابر خرابی سنسورهای کمکی

مقاوم تر می سازد. موقعیت یابی مبتنی بر فیلتر کالمن به شیوه ای کارآمد و دقیق مسئله ردیابی موقعیت ربات را حل می کند. این رویکرد ترکیبی حسگر بهینه به ربات امکان می دهد از مکان شناخته شده اولیه، موقعیت خود را به صورت پیوسته ردیابی کند [۲۰].

با توجه به مرور مطالعات پیشین، چالش های اصلی در ناوبری ربات های متحرک شامل دقت پایین در موقعیت یابی به دلیل نویز حسگرها، عدم قطعیت های محیطی و مسیریابی با کیفیت پایین تولید شده توسط الگوریتم های مسیریابی است. هدف اصلی این مقاله ارائه یک چارچوب ترکیبی مبتنی بر ادغام الگوریتم درخت تصادفی با کاوش سریع (RRT^*) و فیلتر کالمن توسعه یافته (EKF) است که بتواند به صورت یکپارچه به بهبود دقت ناوبری، کاهش خطای موقعیت یابی و افزایش تطبیق پذیری مسیر در شرایط واقعی پردازد. نوآوری مقاله در طراحی یک ساختار حلقه بسته میان موقعیت یابی و مسیریابی است که امکان بازبرنامهریزی مسیر در صورت افزایش خطای تخمین را فراهم می سازد. این چارچوب منجر به تولید مسیریابی هموار، ایمن و قابل اجرا می شود و دقت مسیریابی را نیز بهبود می بخشد. روش پیشنهادی قابلیت کاربرد در ربات های متحرک در حوزه های صنعتی، خدماتی و امدادی را دارد.

ساختار بقیه مقاله به شرح زیر است: در بخش ۲، مسئله موقعیت یابی ربات و روابط مرتبط به طور جامع توضیح داده شده است. در بخش ۳، ترکیب فیلتر کالمن توسعه یافته (EKF) و الگوریتم برنامه ریزی مسیر RRT^* به طور کامل بررسی شده است. در بخش ۴، نتایج پژوهش ارائه شده و در نهایت، بخش ۵ شامل نتیجه گیری کلی و دستاوردهای مقاله می باشد.



شکل ۱- مدل سینماتیک ربات متحرک.

۲- مسئله موقعیت یابی

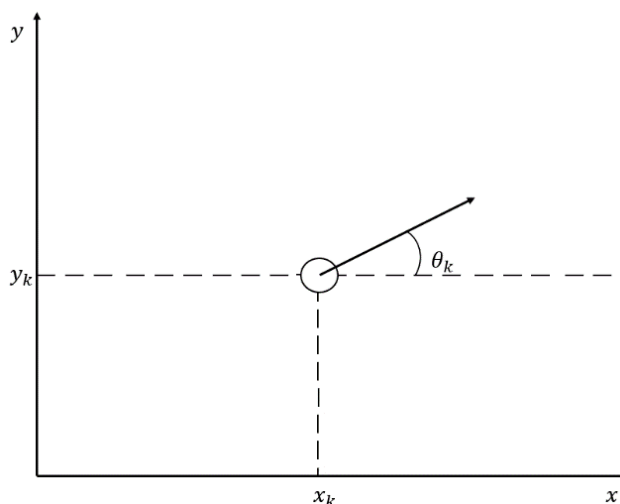
یک ربات متحرک در حین ناوبری در یک محیط شناخته شده با استفاده از ادومتر موقعیت خود را ردیابی می کند. با این حال، عدم قطعیت ادومتری باعث می شود که ربات در مورد صحت موقعیت فعلی خود با عدم قطعیت مواجه باشد. بنابراین، ربات باید خود را نسبت به نقشه محیط موقعیت یابی کند. این امر باعث می شود که عدم قطعیت موقعیت ربات افزایش نیابد. برای مکان یابی، ربات از حسگرهای خارجی خود مانند حسگر لیدار^۱ و حسگر اولتراسونیک^۲ استفاده می کند تا اطلاعاتی درباره محیط خود دریافت کند.

¹ Lidar

² Ultrasonic sensor

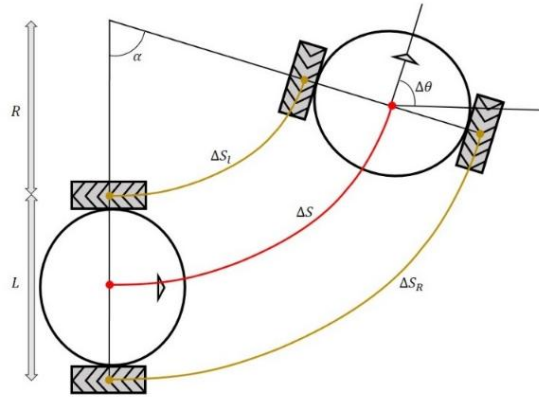
اطلاعات حسگرها را می توان با داده های ادومتری ربات ترکیب کرد تا تعیین موقعیت ربات ممکن شود. اساساً، حتی با وجود حسگر موقعیت یاب جهانی (GPS^1)، موقعیت دقیق ربات نمی تواند به طور مستقیم اندازه گیری شود. سینماتیک یک ربات به پیکربندی چرخ های آن و نوع آن ها بستگی دارد. ساده ترین آن یک پیکربندی دیفرانسیل است. این وسیله نقلیه دارای دو چرخ مستقل است که می توان از آن ها برای کنترل سرعت جلو و زاویه استفاده کرد. در شکل (۱) مدل سینماتیک ربات متحرک نمایش داده شده است. از نظر تئوری، می توان موقعیت واقعی ربات را تا زمانی محاسبه کرد که ساختار ربات به خوبی شناخته شده باشد. سرعت چرخ ربات را می توان به دقت اندازه گیری کرد. از این رو، می توان خطاهای ربات را در طول سفر تشخیص داد و مسیر را بر اساس آن تنظیم نمود. این کار از طریق سینماتیک انجام می شود. ربات باید توانایی تشخیص تغییرات حرکتی هر چرخ را داشته باشد. تخمین موقعیت یک ربات متحرک دشوار است، هیچ راه مستقیمی برای اندازه گیری از حسگرها وجود ندارد و باید حرکت ربات در طول زمان یکپارچه شود همچنین عدم دقت تخمین حرکت (به عنوان مثال، لغزش چرخ) مشکل ساز است و باید یک مدل ربات بر اساس آن تعریف شود. هندسه هر چرخ به حرکت کمک می کند و همچنین محدودیت هایی را اعمال می کند؛ به عنوان مثال، لغزش جانبی در چرخ ها خود به مرور باعث منحرف شدن ربات و در نهایت ایجاد عدم قطعیت در موقعیت یابی می شود. یک ربات متحرک دیفرانسیل را در نظر بگیرید. شکل (۲) نشان دهنده موقعیت ربات با توجه به سیستم مختصات جهانی است. موقعیت ربات با بردار (۱) نشان داده می شود.

$$X_k = \begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} \quad (1)$$



شکل ۲- موقعیت ربات.

¹ Global Positioning System



شکل ۳- یک نمونه چرخش ساده که بیانگر پارامترهای اساسی سینماتیک ربات است.

با توجه به شکل (۱)، حرکت ربات با دو پارامتر اصلی سنجیده می شود: سرعت خطی پیشروی v و سرعت زاویه ای چرخش w . شعاع چرخ های ربات و فاصله بین دو چرخ به ترتیب با R و L نمایش داده می شوند. همچنین زاویه θ بیانگر تغییرات زاویه ای وضعیت ربات نسبت به محور مرجع است. با استناد به شکل (۳)، مسافت طی شده توسط چرخ چپ ΔS_L و چرخ راست ΔS_R داده می شود. همچنین کل مسافت طی شده توسط ربات با ΔS نشان داده شده است. زاویه α ، زاویه بین میله متصل کننده دو چرخ و محور افقی را نشان می دهد. روابط مربوط به محاسبه ΔS_L و ΔS_R و همچنین ΔS به صورت زیر است:

$$\Delta S_L = R\alpha \quad (2)$$

$$\Delta S_R = (R + L)\alpha \quad (3)$$

$$\Delta S = \frac{\Delta S_R + \Delta S_L}{2} \quad (4)$$

با توجه به شکل (۳) مسافت پیموده شده توسط چرخ سمت چپ و چرخ سمت راست قابل محاسبه است. در مرحله بعد زاویه جدید ربات با رابطه (۵) محاسبه می شود:

$$\Delta\theta = \frac{\Delta S_R - \Delta S_L}{L} \quad (5)$$

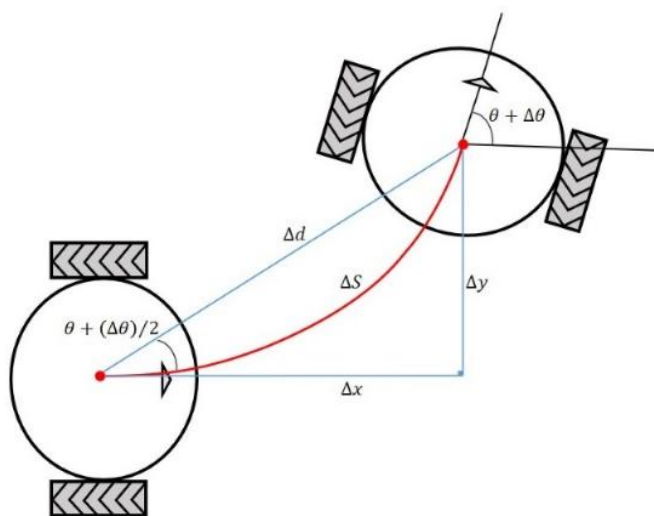
از آنجاییکه از قبل نحوه محاسبه زاویه و مسافت طی شده توسط چرخ های ربات در دسترس است پس امکان محاسبه موقعیت ربات وجود دارد. طبق شکل (۴) بدست آوردن معادلات مربوطه ممکن است.

موقعیت فعلی یک درایو دیفرانسیل را می توان از اطلاعات موقعیت اولیه آن با اضافه کردن فاصله های افزایشی سفر $(\Delta x, \Delta y, \Delta\theta)$ تخمین زد و در نهایت می توان فاصله سنجی ادومتری را به صورت زیر محاسبه کرد:

$$\Delta x = \Delta s \cos\left(\theta + \frac{\Delta\theta}{2}\right) \quad (6)$$

$$\Delta y = \Delta s \sin \left(\theta + \frac{\Delta \theta}{2} \right) \quad (7)$$

وضعیت واقعی ربات تابعی از وضعیت قبلی آن است. در معادله (۸)، فرض بر این است که بردار X_k بیانگر وضعیت فعلی ربات در زمان k است. این وضعیت با استفاده از بردار X_{k-1} ، که نمایانگر وضعیت قبلی ربات در لحظه $k-1$ است، محاسبه می شود. بر این اساس، معادله (۸) نحوه ی گذار از وضعیت پیشین به وضعیت کنونی را توصیف می کند. نکته حائز اهمیت آن است که سیستم مورد بررسی یک سیستم زمان گسسته است، به این معنا که وضعیت ها در گام های زمانی مجزا محاسبه می شوند و مدل سیستم بر مبنای همین گام های گسسته طراحی شده است.



شکل ۴- پارامترهای سینماتیک که مسافت پیموده شده در مسیر انحرافی با آن ها محاسبه می شود.

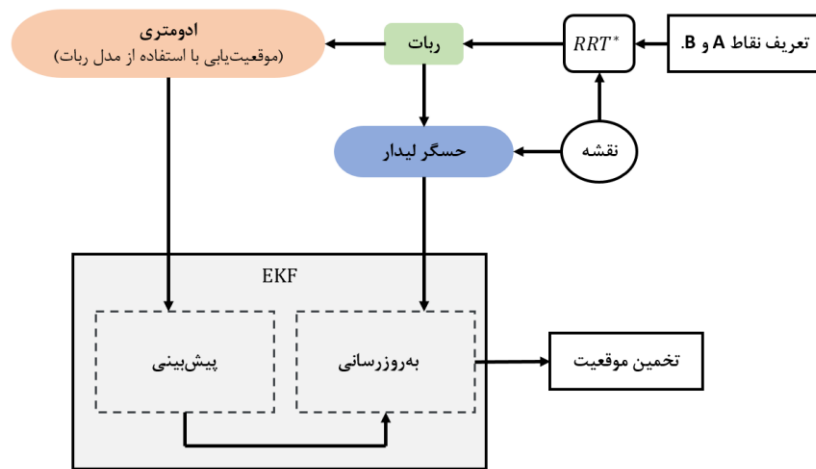
$$X_k = X_{k-1} + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta \theta \end{bmatrix} = X_{k-1} + \begin{bmatrix} \Delta s \cos \left(\theta + \frac{\Delta \theta}{2} \right) \\ \Delta s \sin \left(\theta + \frac{\Delta \theta}{2} \right) \\ \Delta \theta \end{bmatrix} \quad (8)$$

از جمله مهم ترین عوامل ضعف روش ادومتری می توان به انباشته شدن خطاها و بسندگی نداشتن مدل حرکتی فرضی اشاره کرد. ادومتری بر پایه ی اندازه گیری چرخش چرخ ها و محاسبه ی مسافت پیموده شده ی ربات بر اساس این داده ها عمل می کند. با این حال، در طول زمان، خطاهای تصادفی و سینماتیک می توانند در اندازه گیری ها رخ دهند و به صورت تجمعی افزایش یابند، که این امر به تخمین های نادرست از موقعیت و جهت گیری ربات منجر می شود. علاوه بر این، روش ادومتری معمولاً یک مدل حرکت ساده شده را فرض می کند که در آن ربات یا به صورت خطی یکنواخت حرکت می کند یا چرخشی دقیق دارد. در حالی که در عمل، حرکات ربات های متحرک اغلب از رفتارهای غیرخطی محیط تأثیر می پذیرد. این اختلاف میان مدل ایده آل و واقعیت عملی، منشأ بروز خطاهایی در موقعیت یابی بر اساس ادومتری است. بنابراین، برای رفع یا کاهش اثرات این خطاها، بهره گیری از یک رویکرد همجوشی

حسگر نظیر فیلتر کالمن توسعه یافته توصیه می شود تا با ترکیب داده های ادومتری با اطلاعات سایر حسگرها مانند GPS یا لیدار، تخمینی دقیق تر از وضعیت ربات حاصل شود.

۳- موقعیت یابی ربات با EKF و الگوریتم RRT*

چارچوب کلی سیستم یکپارچه EKF-RRT* در شکل (۵) نشان داده شده است. فرآیند با تعریف نقاط A و B به عنوان نقطه شروع و هدف آغاز می شود و الگوریتم RRT* بهترین مسیر ممکن را برای ربات محاسبه می کند. حسگرهای داخلی ربات، شامل شمارنده های چرخ های چپ و راست، در کنار مدل حرکتی ربات برای تخمین اولیه موقعیت به کار می روند. همچنین، حسگر لیدار به عنوان یک حسگر خارجی نقشه محور، نقش کلیدی در اصلاح موقعیت نسبی ربات در مرحله به روزرسانی دارد. در این ساختار، از فیلتر کالمن توسعه یافته (EKF) در هر دو مرحله پیش بینی و تصحیح استفاده می شود؛ به طوری که مرحله پیش بینی مبتنی بر داده های ادومتری و مرحله تصحیح بر اساس اندازه گیری های لیدار انجام می گیرد. نکته مهم آن است که دقت موقعیت یابی ربات که توسط EKF تأمین می شود، نقش اساسی در امکان پذیر بودن و پایداری مسیر پیشنهادی RRT* دارد. مسیر بهینه محاسبه شده توسط RRT* بدون تخمین صحیح موقعیت ربات ممکن است در عمل غیرقابل پیاده سازی باشد. از سوی دیگر، هرگونه انحراف یا افزایش عدم قطعیت که توسط EKF شناسایی شود، می تواند منجر به فراخوانی مجدد RRT* و بازبرنامه ریزی مسیر شود. این تعامل دوطرفه، منجر به بهبود عملکرد ناوبری در شرایط غیرایستا و نامطمئن می شود.



شکل ۵- چهارچوب کلی رویکرد پیشنهادی.

۳-۱- موقعیت یابی ربات مبتنی بر EKF

EKF یک نسخه غیرخطی از فیلتر کالمن است که برای تخمین وضعیت یک سیستم دینامیکی غیرخطی با نویز گاوسی طراحی شده است. EKF از مدل ادومتری برای تخمین موقعیت ربات استفاده می کند. همچنین به کمک این رویکرد می توان اطلاعات حسگرهای داخلی و خارجی را ترکیب کرد. فرض کنید که موقعیت ربات متحرک در زمان $k - 1$ به صورت زیر است:

$$X_{k-1} = \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \theta_{k-1} \end{bmatrix} \quad (9)$$

در لحظه k ، ربات به سمت موقعیت جدید X_k حرکت می کند. با استفاده از معادله (۱۰)، موقعیت جدید X_k در زمان k می تواند از موقعیت قبلی X_{k-1} پیش بینی شود:

$$X_k = f(X_{k-1}, u_k) = \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \theta_{k-1} \end{bmatrix} + \begin{bmatrix} \frac{\Delta s_r + \Delta s_l}{2} \cos\left(\theta_{k-1} + \frac{\Delta s_r - \Delta s_l}{2L}\right) \\ \frac{\Delta s_r + \Delta s_l}{2} \sin\left(\theta_{k-1} + \frac{\Delta s_r - \Delta s_l}{2L}\right) \\ \frac{\Delta s_r - \Delta s_l}{L} \end{bmatrix} \quad (10)$$

در رابطه (۱۰)، k زمان نمونه برداری و Δs_r جابجایی چرخ راست و Δs_l جابجایی چرخ چپ است. همچنین x_{k-1} و y_{k-1} موقعیت ربات در زمان $k-1$ و θ_{k-1} جهت ربات در زمان $k-1$ است. L طول محور ربات که دو چرخ چپ و راست را به هم متصل می کند، می باشد. اجزای ورودی کنترل u_k مطابق (۱۱) است:

$$u_k = \begin{bmatrix} \Delta s_l \\ \Delta s_r \end{bmatrix} \quad (11)$$

EKF به کمک بسط تیلور، دینامیک سیستم و مدل های اندازه گیری را خطی سازی می کند تا بتواند تخمین وضعیت و کوواریانس را پیش بینی و به روزرسانی نماید. در این رویکرد، دینامیک سیستم غیر خطی به صورت زیر مدل می شود:

$$X_k = f(X_{k-1}, u_{k-1}) + w_{k-1} \quad (12)$$

$$Z_k = h(X_k) + v_k \quad (13)$$

X_k بردار متغیرهای حالت در نمونه k است، همچنین w_k و v_k به ترتیب نویز فرآیند و نویز اندازه گیری در لحظه k هستند و خروجی اندازه گیری در همان لحظه است. با توجه به رابطه (۱۰)، تابع غیر خطی f را می توان به صورت رابطه (۱۴) بیان کرد:

$$f(X_{k-1}, u_{k-1}) = \begin{bmatrix} x_{k-1} + \Delta s \cos(\Delta\theta) \\ y_{k-1} + \Delta s \sin(\Delta\theta) \\ \theta_{k-1} + \Delta\theta \end{bmatrix} \quad (14)$$

رابطه (۱۵) ژاکوبین تابع غیر خطی سیستم است:

$$F = \begin{bmatrix} 1 & 0 & -\Delta s \sin(\theta_k + \Delta\theta) \\ 0 & 1 & \Delta s \cos(\theta_k + \Delta\theta) \\ 0 & 0 & 1 \end{bmatrix} \quad (15)$$

h تابع غیر خطی اندازه گیری سیستم و u_k سیگنال ورودی سیستم در نمونه k است. واضح است که تابع f با مدل واقعی و عملی سیستم اختلاف دارد. این اختلاف در نویز فرآیند مدل سازی خواهد شد، که ناشی از عدم قطعیت در مدل سازی ریاضی سیستم است. نویز اندازه گیری همان نویز سنسورهاست که برخی از متغیرهای حالت و یا تابعی از آن ها را اندازه گیری می کند.

نویز اندازه گیری و نویز فرآیند اغلب به صورت گاوسی مدل سازی می شوند و ماتریس کوواریانس آن ها به ترتیب با Q_k و R_k نمایش داده می شوند. با انجام تحلیل های آماری بر روی سنسورها می توان R_k را به دقت تعیین کرد؛ اما به دلیل ماهیت پیچیده و متنوع سیستم های غیر خطی، روش مشخص و قطعی برای تعیین R_k و Q_k وجود ندارد و این ماتریس ها معمولاً به صورت سعی و خطا تنظیم می شوند. الگوریتم EKF در دو فاز پیش بینی و به روز رسانی اجرا می شود. فرض کنید که تمام اطلاعات سیستم تا نمونه $k - 1$ در دسترس باشد. در این صورت، بر اساس مدل ریاضی سیستم و ورودی های موجود تا لحظه $k - 1$ ، تخمینی اولیه از متغیرهای حالت در لحظه k محاسبه می شود. این مرحله، فاز پیش بینی EKF نام دارد. در لحظه k ، اندازه گیری جدیدی توسط سنسورها انجام شده و در دسترس قرار می گیرد. در مرحله دوم، یعنی فاز به روز رسانی، تخمین اولیه بر اساس داده های جدید اصلاح می شود تا مقدار نهایی متغیرهای حالت در زمان k به دست آید. فاز پیش بینی توسط دو رابطه (۱۶) و (۱۷) انجام می شود:

$$\hat{X}_{k|k-1} = f(\hat{X}_{k-1}, u_{k-1}) \quad (16)$$

$$P_{k|k-1} = F_{k-1}P_{k-1}F_{k-1}^T + Q_{k-1} \quad (17)$$

در روابط بالا \hat{X}_{k-1} بردار حالت های تخمین در نمونه $k - 1$ و P_{k-1} ماتریس کواریانس در نمونه $k - 1$ و $\hat{X}_{k|k-1}$ پیش بینی متغیر حالت در نمونه k بر اساس اطلاعات سیستم تا نمونه $k - 1$ و $P_{k|k-1}$ نیز پیش بینی ماتریس کواریانس در لحظه k است. ماتریس F_k دینامیک حالت خطی شده است که از طریق خطی سازی تیلور به صورت زیر به دست می آید:

$$F_k = \left. \frac{\partial f}{\partial X} \right|_{\hat{X}_{k-1}} \quad (18)$$

H_k نیز دینامیک اندازه گیری خطی شده است که در فاز به روز رسانی استفاده می شود. این ماتریس نیز به صورت زیر محاسبه می شود:

$$H_k = \left. \frac{\partial h}{\partial X} \right|_{\hat{X}_{k-1}} \quad (19)$$

فاز به روز رسانی با استفاده از روابطی که در ادامه ارائه می شوند انجام می گیرد. این مرحله، از اطلاعات حسگر برای اصلاح و بهبود تخمین حالت به دست آمده در فاز پیش بینی بهره می برد. در این مقاله، سنسور یک آشکارساز لیدار است، نشانه ها در محیط عملیاتی مرتب شده اند. ربات از قبل موقعیت هر یک را می داند، هر نشانه یک شناسه منحصر به فرد دارد. بردار اندازه گیری سنسور آشکارساز شامل سه پارامتر r و α و \hat{t} است که هر کدام به ترتیب بیانگر فاصله از ربات تا لیدار، زاویه از ربات نسبت به لیدار و شناسه منحصر به فرد مربوط به نشانه هستند. مشاهدات حسگر با Z نشان داده می شود:

$$Z_k = \begin{bmatrix} r_k \\ \alpha_k \end{bmatrix} \quad (20)$$

خوانش حسگر لیدار، همراه با نویز اندازه گیری v با کوارینانس R است. خوانش های مورد انتظار حسگر با \hat{Z} نمایش داده می شود و از طریق رابطه (۲۱) قابل محاسبه است:

$$\hat{Z}_k = h(\hat{X}_{k|k-1}, S_i) \quad (21)$$

در رابطه (۲۱)، S_i یک بردار حاوی مختصات نقاط عطف و شناسه مخصوص آن ها است که این مختصات از قبل تعریف شده اند:

$$S_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (22)$$

در بردار (۲۲)، x_i و y_i مختصات مربوط به نشانه و \hat{l} شناسه مخصوص آن نشانه است. همچنین در (۱۷)، $\hat{X}_{k|k-1}$ بردار حالت بدست آمده از مرحله پیش بینی است. h تابعی است که خوانش مورد انتظار حسگر را با توجه به بردار حالت و بردار لیدار محاسبه می کند. در (۲۳)، تابع غیرخطی h آمده است:

$$h(\hat{X}_{k|k-1}, S_i) = \begin{bmatrix} l \\ \beta \end{bmatrix} = \begin{bmatrix} \sqrt{(S_{i,x} - x)^2 + (S_{i,y} - y)^2} \\ \text{atan2}(S_{i,x} - x, S_{i,y} - y) - \theta \end{bmatrix} \quad (23)$$

در (۲۴) ژاکوبین تابع غیرخطی h است:

$$H = \begin{bmatrix} \frac{-(S_{i,x} - x)}{\sqrt{(S_{i,x} - x)^2 + (S_{i,y} - y)^2}} & \frac{-(S_{i,y} - y)}{\sqrt{(S_{i,x} - x)^2 + (S_{i,y} - y)^2}} & 0 \\ \frac{(S_{i,y} - y)}{\sqrt{(S_{i,x} - x)^2 + (S_{i,y} - y)^2}} & \frac{-(S_{i,x} - x)}{\sqrt{(S_{i,x} - x)^2 + (S_{i,y} - y)^2}} & -1 \end{bmatrix} \quad (24)$$

در رابطه (۲۳)، با توجه به اینکه ربات مجهز به فاصله یاب لیزری است، این حسگر فاصله سمت l و زاویه β را اندازه گیری می کند. \hat{l} شناسه مخصوص مربوط به نشانه است. در رابطه (۲۴)، تابع اندازه گیری خطی سازی شده است. با رسیدن اندازه گیری های حسگر، فاز به روزرسانی اجرا می شود:

$$K = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \quad (25)$$

$$\hat{X}_k = \hat{X}_{k|k-1} + K_k (Z_k - \hat{Z}_k) \quad (26)$$

$$P_k = (I - K_k H_k) P_{k|k-1} \quad (27)$$

در فرمول (۱۹)، Z_k همان خروجی اندازه گیری شده توسط سنسورها است. بنابراین، در قدم اول بهره کالمن (K_k) محاسبه می شود. بهره کالمن مقداری است که میزان ادغام اندازه گیری در حالت جدید را مشخص می کند. سپس تخمین نهایی متغیرهای حالت \hat{X}_k و ماتریس کواریانس P_k در نمونه k محاسبه خواهد شد. اگر به عبارت به روزرسانی حالت ها دقت کنیم، متوجه خواهیم شد که در فاز به روزرسانی به حالت پیش بینی شده در گام اول یک ترم جدید اضافه شده است که متناسب با خطای پیش بینی است. عبارت $(Z_k - \hat{Z}_k)$ دقیقاً بیانگر اختلاف بین خروجی واقعی اندازه گیری شده توسط سنسورها و خروجی به دست آمده در فاز پیش بینی است.

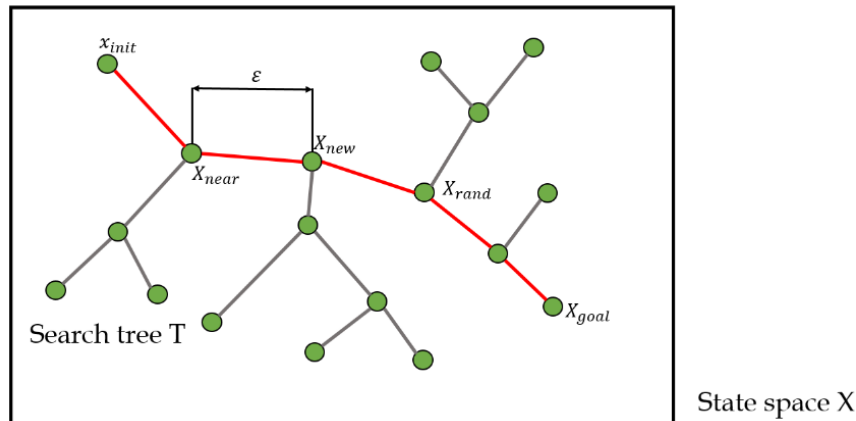
۳-۲- الگوریتم مسیریابی RRT*

الگوریتم RRT* یک روش برنامه ریزی مسیر مبتنی بر نمونه برداری است که با هدف یافتن مسیرهای مؤثر و امکان پذیر برای سیستم های رباتیکی در محیط های پیچیده توسعه یافته است [۲۲]. این الگوریتم با رشد تدریجی یک ساختار درختی از یک پیکربندی اولیه به سمت پیکربندی هدف، فضای حالت را به صورت هوشمندانه برای یافتن مسیرهای بهینه کاوش می کند [۲۳]. یکی از ویژگی های کلیدی RRT*، قابلیت بهینه سازی پویا ساختار درختی از طریق باز تنظیم گره ها است که موجب بهبود کیفیت مسیر و همگرایی آن به مسیر بهینه می شود [۲۴]. این ویژگی تطبیقی به RRT* اجازه می دهد تا فضاهای با ابعاد بالا و سیستم های دارای محدودیت های غیر هولونومیک را به طور مؤثر مدیریت کند [۲۵]. برای شروع، برخی از تعاریف و رویه های اولیه RRT* شرح داده شده است. پیکربندی ناحیه ناشناس به صورت X نشان داده می شود، جایی که X یک مجموعه با اعضای $X = \{x, y, \theta\}$ است. در اینجا x و y مختصات دو بعدی را نشان می دهند، در حالی که θ نشان دهنده جهت است. علاوه بر این، تمام مناطق مربوط به موانع را در بر می گیرد، در حالی که X_{free} مربوط به قسمت هایی از محیط است که بدون مانع هستند. رابطه زیر بیانگر ارتباط ریاضی بین این دو مجموعه است:

$$X_{free} = \frac{X}{X_{obs}} \quad (28)$$

در ادامه مفاهیم مهم تری برای بررسی مطرح می شوند این مفاهیم عبارت اند از: x_{init} و x_{goal} که به ترتیب به معنی نقطه شروع و نقطه هدف هستند [۲۶]. راه حل مسئله برنامه ریزی حرکت مسیر $\sigma: [0,1] \rightarrow X_{free}$ که $\sigma(0) = x_{init}$ و $\sigma(1) = x_{goal}$ است. در RRT*، همچنین یک الگوی درختی $T = (V, E)$ ساخته می شود که از مجموعه رئوس $V \subset X$ و شاخه های $E \subset V \times V$ تشکیل شده است. نمونه به طور تصادفی یک حالت $x_{rand} \in X_{free}$ را از فضای بدون مانع نمونه برداری می کند. فاصله با دو حالت $x, x' \in X$ ، تابع $dist(x, x')$ هزینه مسیر اتصال x و x' را برمی گرداند. تابع هزینه استفاده شده در این مقاله بر حسب فاصله اقلیدسی است. $Nearest$ و $Near$ با یک مجموعه $V \subset X$ و حالت $x \in X$ تعریف می شوند. این دو رویه نزدیکترین گره در درخت و تمام گره های نزدیک را که درون یک توپ با حجم $\gamma ((\log n)/n)^{1/d}$ در مرکز قرار دارند، برمی گرداند. در حدود x از نظر فاصله γ ثابتی است که در [۲۶] توضیح داده شده است، d بعد فضای X و n تعداد رئوس است. هدایت تابع $Steer(x, x')$ یک مسیر $\sigma: [0,1] \rightarrow X$ را برمی گرداند که x و x' را به هم متصل می کند. بررسی برخورد با یک مسیر $\sigma: [0,1] \rightarrow X$ تابع $InCollision(\sigma)$ بررسی می کند که آیا مسیر در X_{free} قرار دارد یا نه، و اگر نه، اولین حالت

شکست خورده x_{fail} را برمی گرداند. تابع هزینه با یک راس x ، اجازه می دهیم $C(x)$ هزینه کل مسیر منحصر به فرد باشد که از راس ریشه x_{init} شروع می شود، در امتداد گره های درخت می رود و به حالت x می رسد، در حالی که $C(\sigma)$ تابع هزینه از مسیر $\sigma: [0,1] \rightarrow X$ است. مرتب سازی لیست L ، مجموعه ای مرتب از عناصر با شکل (c_i, x_i, σ_i) ، تابع $L.sort()$ همه عناصر L را بر اساس c_i به ترتیب صعودی مرتب می کند.



شکل ۶- توسعه برنامه ریزی مسیر RRT*.

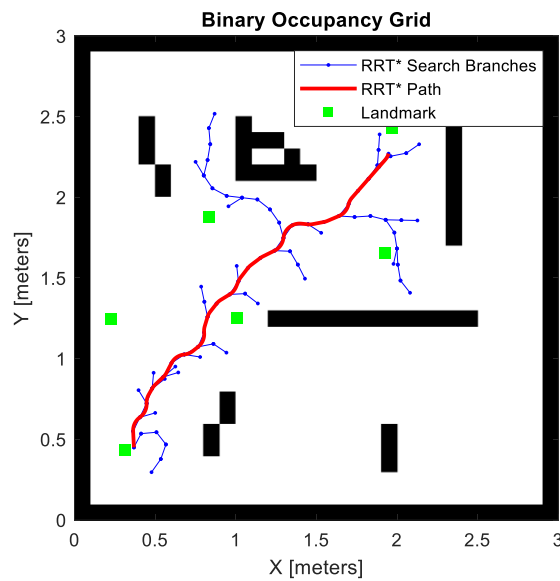
در پیکربندی، RRT* درخت جستجو را بر اساس گره ها با جمع آوری تصادفی نمونه ها به صورت تدریجی ایجاد می کند، بنابراین گودها به گونه ای ساخته می شوند که به سرعت فاصله مورد انتظار گره های انتخابی تصادفی در درخت را کاهش می یابد. همانطور که گره ها رشد می کنند، با استفاده از x_{rand} (گره های تصادفی)، برای اتصال گره x_{near} و گره x_{rand} تولید می شود [۲۶]. در شکل (۶) نشان داده شده است که با تکرار این فرآیند، گره ها با هم یک درخت را تشکیل می دهند تا به حالت هدف برسند. شبه کد RRT* بصورت الگوریتم (۱) است.

الگوریتم ۱- شبه کد RRT*.

[1] Algorithm 1: The RRT* Algorithm

- [2] 1: $V \leftarrow \{x_{init}\}; E \leftarrow \emptyset; T \leftarrow (V, E);$
- [3] 2: **for** $i = 1 \rightarrow N$ **do**
- [4] 3: $x_{rand} \leftarrow Sample(i);$
- [5] 4: $x_{near} \leftarrow Near(V, x_{rand});$
- [6] 5: **if** $x_{near} = \emptyset$ **then**
- [7] 6: $x_{near} \leftarrow Nearst(V, x_{rand});$
- [8] 7: **end if**
- [9] 8: $x_{parent} \leftarrow Find\ Best\ Parents(x_{near}, x_{rand});$
- [10] 9: **if** $x_{parent} \neq Null$ **then**
- [11] 10: $V \leftarrow V \cup \{x_{rand}\}; E \leftarrow E \cup \{(x_{near}, x_{rand})\};$
- [12] 11: $E \leftarrow Rewire(E, x_{near}, x_{rand});$
- [13] 12: **end if**
- [14] 13: **end for**
- [15] 14: **return** $T = (V, E);$

الگوریتم RRT^* مبتنی بر نمونه گیری و غیرمتمرکز است و فرآیند هدر رفتن در حضور تعداد زیادی مجموعه داده برای پردازنده را کاهش می دهد. علاوه بر این، در حین گشت زنی و کاوش، RRT^* مسیری بهینه برای دور زدن موانع پیدا می کند. این بدان معناست که ربات می تواند با استفاده از اجتناب از موانع و مسیریابی با تولید مدل های احتمالی محیط، در یک منطقه ناشناخته حرکت کند. RRT^* با در نظر گرفتن مدل ربات به واقعیت نزدیک است، بدون اینکه فرض کنیم این ربات دارای قابلیت حرکت در همه جهات است. بر اساس مدل های سینماتیکی، مشخص است که هر تعداد از موقعیت ها را می توان پس از مدت زمان مشخصی توسط ربات به دست آورد و به همین ترتیب، تعداد موقعیت جدید بر اساس موقعیت قبلی ایجاد می شود. این باعث ایجاد تعداد زیادی از مقادیر ممکن در طول زمان می شود. بر اساس هر موقعیت x_{rand} مورد انتظار، نزدیک ترین موقعیت x_{near} را همیشه می توان در تعداد زیادی از مقادیر یافت. سپس مجموعه ای از مسیرهای متناظر با معکوس کردن نقاط وضعیت به دست می آید. روش مورد استفاده در اینجا مشابه استراتژی نمونه گیری تصادفی است و سپس بهترین ورودی کنترلی از نتایج نمونه انتخاب می شود. همچنین مواجهه با مشکلات مربوط به موانع بسیار ساده است: زمانی که x_{rand} در حال برخورد با مانع است، مستقیماً نادیده گرفته می شود و اگر مسیر مانع را قطع کند، x_{near} دیگری مستقیماً انتخاب می شود. این به مجموعه ای از مسیرهای مختلف برای اجتناب از برخورد با موانع منجر می شود. سپس بر این اساس یک طرح مسیر تولید می شود و می توان نتایج را ترسیم و بررسی کرد.



شکل ۷- مسیر نامی طراحی شده توسط الگوریتم RRT^* و نقشه محیط.

۴-نتایج

ربات در محیط نقشه حرکت می کند و اطلاعات اندازه گیری را از سنسورهای داخلی حرکت چرخ های خود ثبت می کند و با حسگرهای خارجی خود اندازه گیری های مربوط به نشانه های محیط و موانع را جمع آوری می کند. محیط شبیه سازی مربوطه در متلب شبیه سازی شده است. برای حل مشکل دقت ناوبری، روش موقعیت یابی مبتنی بر EKF و روش برنامه ریزی مسیر مبتنی بر RRT^* به ترتیب معرفی شده است. محیط شبیه سازی ربات به نوعی شبیه به یک محیط بسته فرضی طراحی شده است. در شکل (۷) مسیر واقعی ربات و نقشه محیط نشان داده شده است. مسیر حرکت ربات با رنگ قرمز مشخص شده است. مسیر حرکت ربات توسط

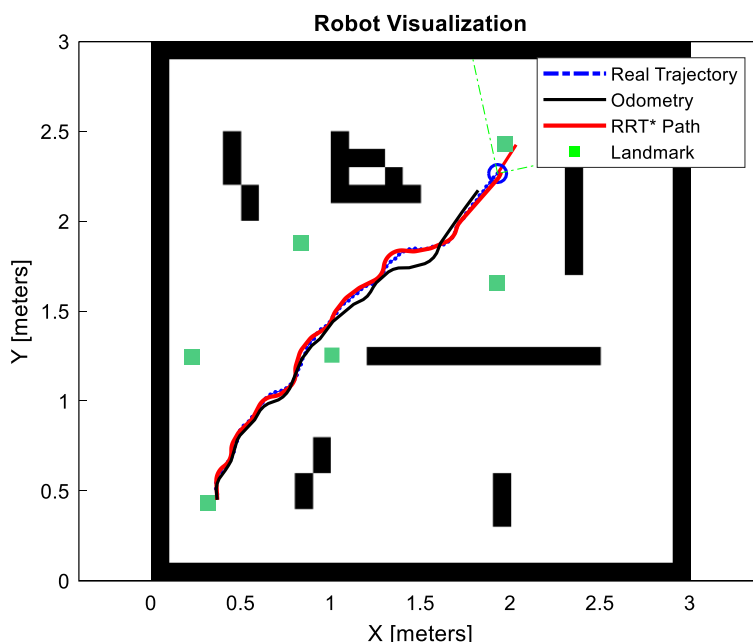
الگوریتم درختان تصادفی با کاوش سریع طراحی می شود. موقعیت یابی با شناسایی نشانه ها از داده های اندازه گیری شده و مطابق ساختن آن ها با نشانه های موجود در نقشه انجام می شود. از تفاوت بین موقعیت نشانه های تخمین زده شده و اندازه گیری شده، برای محاسبه موقعیت و وضعیت ربات استفاده می شود. این ربات مجهز به حسگر لیدار است. حسگر لیدار در جلوی ربات نصب شده است و فاصله و زاویه نسبت به نشانه مشاهده شده را اندازه گیری می کند. سرعت ربات 1 m/s است و ماکزیمم زاویه فرمان آن $1/59$ درجه در نظر گرفته شده است. فاصله بین چرخ ها $0/05$ و ماکزیمم دامنه فاصله یاب لیزری $0/5$ متر و میدان دید آن 90 درجه است. فرکانس کنترل 50 Hz و انحراف استاندارد نویز سرعت مستقیم و سرعت زاویه ای به ترتیب σ_v و σ_γ بصورت زیر است:

$$\sigma_v = 0.01 \text{ m/s}, \sigma_\gamma = 0.017^\circ$$

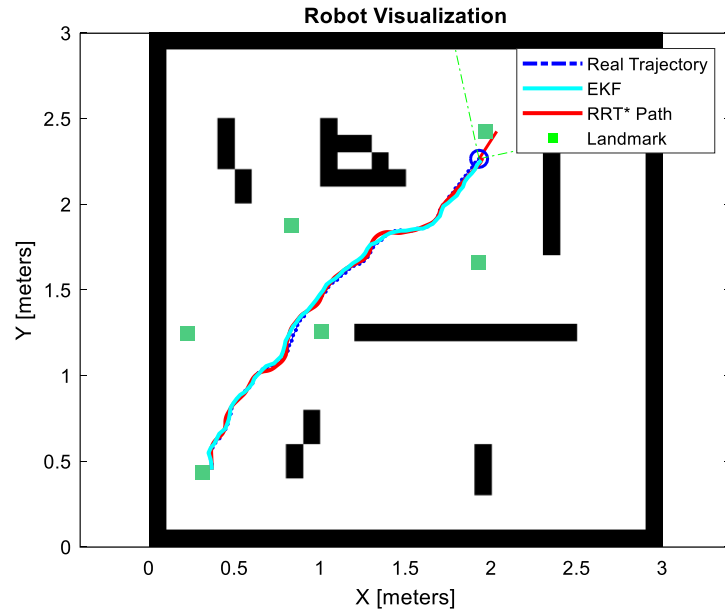
فرکانس نمونه برداری سنسورها 5 Hz و انحراف استاندارد نویز اندازه گیری در فاصله و سمت به ترتیب σ_r و σ_θ بصورت زیر است:

$$\sigma_r = 0.5 \text{ m/s}, \sigma_\theta = 0.05^\circ$$

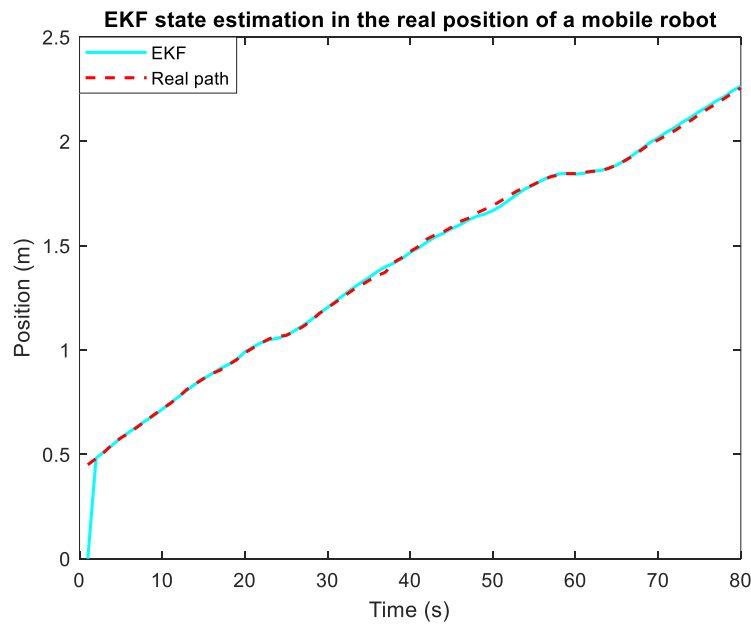
شکل های (۸) و (۹)، نتایج حاصل از موقعیت یابی بر اساس ادومتری و فیلتر کالمن توسعه یافته را نشان می دهند. نشانه های موجود در نقشه با علامت ■ مشخص شده است. در شکل (۹) خط مشکی مربوط به موقعیت یابی ربات توسط ادومتری است. در شکل (۹) رنگ فیروزه ای مربوط به تخمین EKF است. کاملاً مشخص است که نتایج عملکرد روش فیلتر کالمن توسعه یافته بهتر از ادومتری است. به عبارت دیگر، در موقعیت یابی با الگوریتم پیشنهادی، مسیر تخمین زده شده وسیله تا حد امکان نزدیک به مسیر واقعی می باشد. ربات از نقطه شروع تا نقطه هدف را با موفقیت موقعیت یابی کرده و طبق برنامه ریزی مسیر که توسط الگوریتم RRT^* طراحی شده، پیش رفته و به نقطه هدف رسیده است. در ترکیب الگوریتم EKF و RRT^* مشخص است که مسیر تولید شده توسط شبیه سازی الگوریتم پیشنهادی بهینه است. در شکل (۱۰) دقت موقعیت یابی EKF آمده است.



شکل ۸- مسیر واقعی و تخمین موقعیت ربات توسط ادومتری.

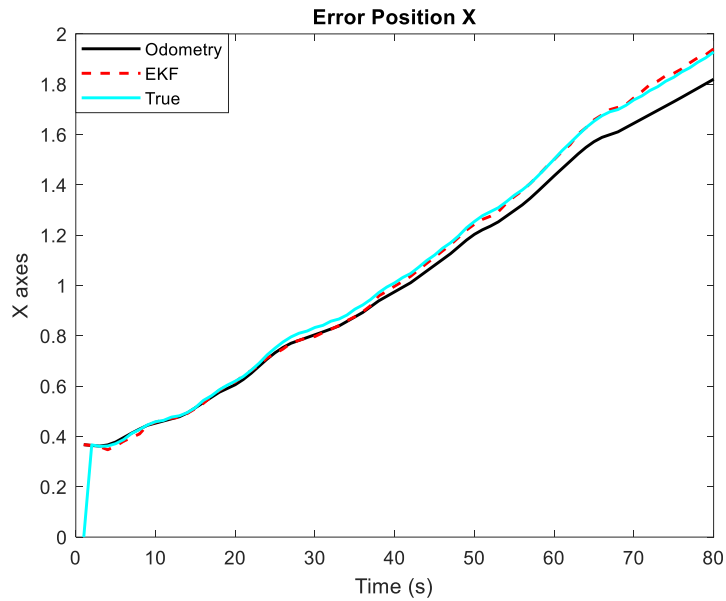


شکل ۹- مسیر واقعی و تخمین موقعیت ربات توسط EKF.



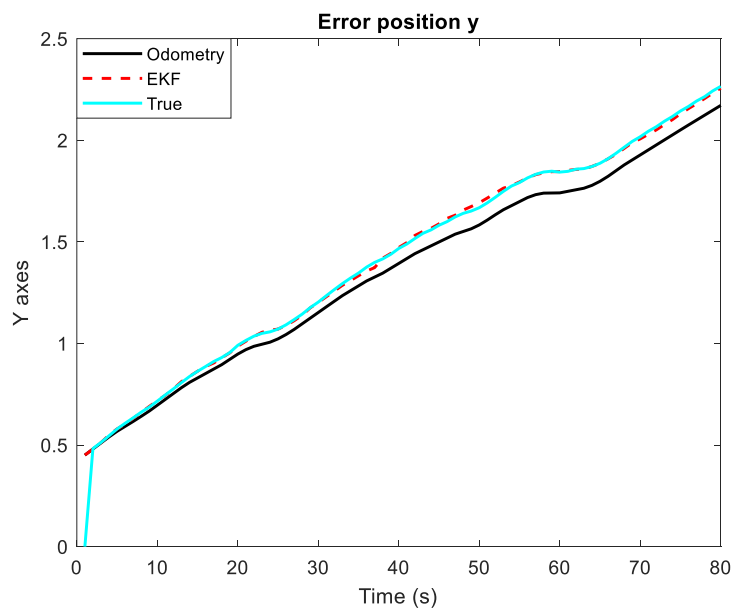
شکل ۱۰- دقت تخمین EKF در رابطه با تخمین موقعیت حقیقی ربات متحرک.

در شکل (۹)، خطای موقعیت یابی و خطای زاویه ایجاد شده بسیار کوچک است و خطا افزایشی نیست. در نتیجه، رویکرد پیشنهادی موجب افزایش دقت موقعیت یابی به طور قابل توجهی در مقایسه با ادمتری شده است. در شکل های (۱۱) تا (۱۳) خطای موقعیت یابی در محورهای X و Y و همچنین جهت گیری زاویه جلوی ربات نسبت به زاویه مسیر آمده است.

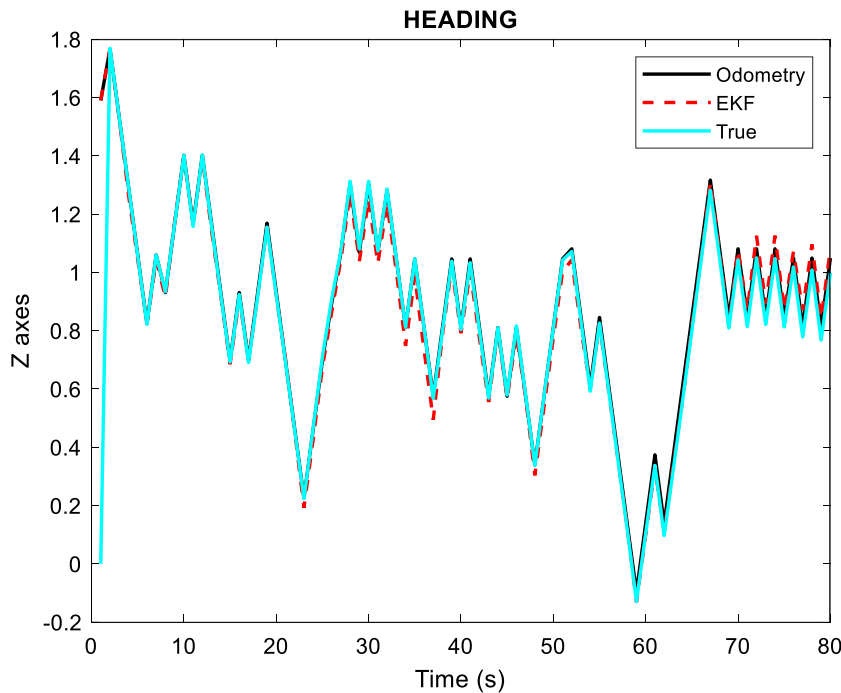


شکل ۱۱- دقت تخمین EKF در مقایسه با محاسبات ادومتری در محور X .

در شکل (۱۱) خطای موقعیت یابی EKF با ادومتری در محور X مقایسه شده است ادومتری تنها در همان لحظات ابتدایی مطابق موقعیت واقعی ربات عمل کرده و انباشته شدن نویزها باعث عملکرد ناموفق ادومتری شده است این به دلیل در نظر نگرفتن نویز سنسورها و نویزهای مربوط به سرعت زاویه ای و سرعت پیشروی است. اما، در خصوص موقعیت یابی با EKF به دلیل اینکه مشخصات نویزها در نظر گرفته شده اند و اطلاعات حسگرهای درونی و بیرونی با هم ترکیب شدند، دیگر ربات با مشکل انباشته شدن خطاها مواجه نمی شود. در شکل (۱۲) خطای موقعیت یابی EKF و ادومتری در محور Y مقایسه شده است. خطاهای موقعیت یابی مانند خطای سیستم و خطای حسگرها که با استفاده از الگوریتم ادومتری ایجاد می شود، به طور پیوسته در حال جمع شدن هستند.



شکل ۱۲- دقت تخمین EKF در مقایسه با محاسبات ادومتری در محور Y .



شکل ۱۳- دقت تخمین EKF در مقایسه با محاسبات ادومتری مربوط به زاویه دیفرانسیل ربات در جهت گیری.

در شکل (۱۳) جهت گیری زاویه جلوی ربات تقریباً به درستی زاویه مسیر را دنبال کرده است. جذر مربع خطا^۱ (RMSE) معیاری است که معمولاً برای اندازه گیری تفاوت بین مقادیر پیش بینی شده و مقادیر واقعی در تحلیل رگرسیون استفاده می شود. RMSE میانگین انحراف یا خطا بین مقادیر پیش بینی شده و واقعی پیش بینی های یک مدل را نشان می دهد. با گرفتن جذر میانگین مجذور خطاهای بین مقادیر پیش بینی شده و مقادیر واقعی محاسبه می شود. مقدار RMSE کمتر نشان دهنده برازش بهتر مدل است. به منظور ارزیابی جامع تر کارایی روش پیشنهادی، عملکرد آن با دو روش مرجع مبتنی بر فیلتر کالمن توسعه یافته و یک روش مبتنی بر ادومتری، مطابق با مطالعات اخیر مقایسه شده است. نتایج این مقایسه در جدول X ارائه شده است. همان گونه که مشاهده می شود، روش پیشنهادی، در تمامی شاخص های کلیدی شامل میانگین خطای جهت X و Y، خطای زاویه دیفرانسیل و مقدار RMSE، عملکرد بهتری نسبت به سایر روش ها از خود نشان داده است. همچنین زمان محاسباتی این روش قابل رقابت با روش های رقیب است. این نتایج نشان دهنده دقت بالاتر و پایداری بیشتر روش پیشنهادی در ردیابی و ناوبری ربات متحرک است. به طور خاص، مقدار RMSE به دست آمده برای روش پیشنهادی ۰/۰۵۵ متر است که نسبت به روش های به ترتیب بهبود قابل توجهی را نشان می دهد. همچنین خطای زاویه و میانگین خطا در هر دو راستا نیز در روش پیشنهادی کمتر از سایر روش ها است و این موضوع بیانگر افزایش دقت تخمین موقعیت و جهت یابی ربات است. در مجموع، نتایج جدول X تأیید می کند که ترکیب فیلتر کالمن توسعه یافته با الگوریتم پیشنهادی می تواند منجر به بهبود معنی دار عملکرد ردیابی و ناوبری در سناریوهای عملیاتی شود.

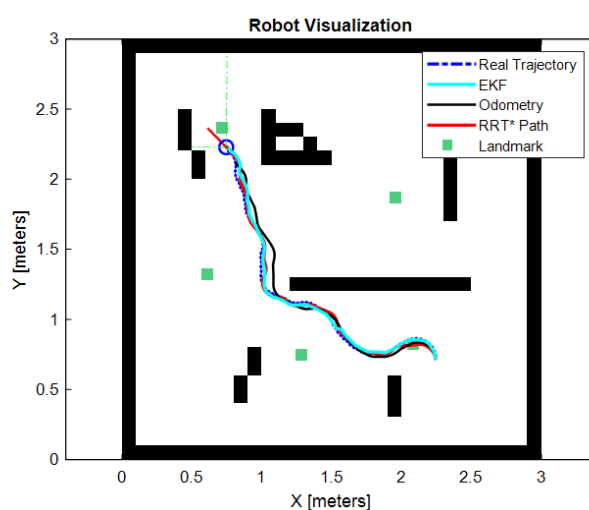
¹ Root Mean Square Error

جدول ۱- مقایسه عملکرد روش پیشنهادی با سایر روشها

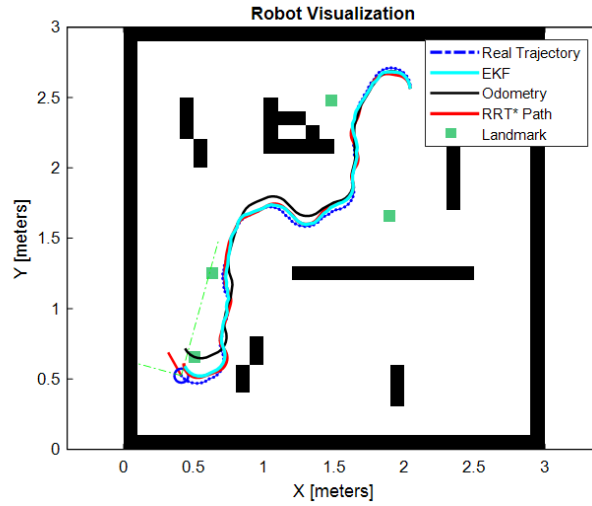
استراتژی ناوبری [۲۰]	[۱۹] میانگین خطا (متر)		[۱۸] خطای زاویه دیفرانسیل (درجه)	RMSE [۱۷]	[۱۶] زمان محاسبات (ثانیه)
	[۲۲] درجهت X	[۲۱] ذر جهت Y			
EKF + RRT* [۲۸]	۰/۰۲۷ [۲۷]	۰/۰۲۵ [۲۶]	۰/۰۲۵ [۲۵]	۰/۰۵۵ [۲۴]	۱/۵۷ [۲۳]
IEKF [۲۸]	۰/۰۸ [۳۳]	۰/۰۷ [۳۲]	۰/۰۴ [۳۱]	۰/۰۶۹ [۳۰]	۱/۷ [۲۹]
EKF [۲۷]	۰/۰۹۹ [۳۹]	۰/۰۹ [۳۸]	۰/۰۴۵ [۳۷]	۰/۰۷۴ [۳۶]	۱/۵ [۳۵]
Odometry [۴۶]	۰/۱۱۵ [۴۵]	۰/۱۰۲ [۴۴]	۰/۰۵۵ [۴۳]	۰/۰۸۵ [۴۲]	۱/۳۴ [۴۱]

۴-۱- بررسی عملکرد در مسیرهای مختلف

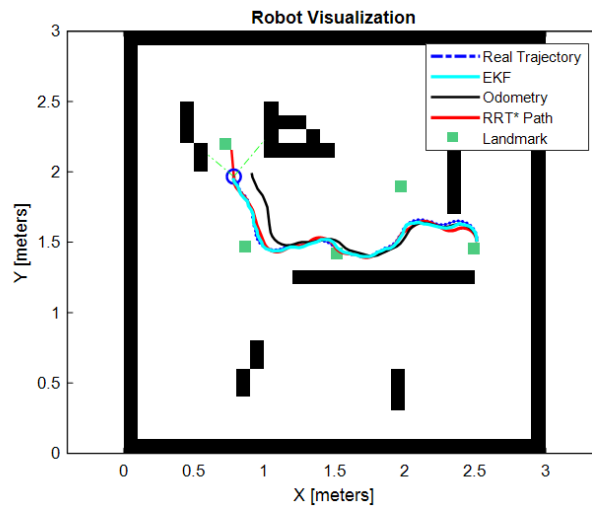
برای بررسی بهتر الگوریتم برنامه ریزی مسیر، عملکرد RRT* در حالت های مختلف آزمایش شد. شبیه سازی در چهار گروه مختلف از مسیرهای انتخابی انجام می شود. ربات متحرک از نقطه شروع تا نقطه هدف در پایین حرکت می کند. مسیرهای قرمز نشان دهنده نتایج الگوریتم RRT* است. مسیرهای سیاه نشان دهنده نتایج بر اساس محاسبات اودومتري هستند. مسیرهای فیروزه ای نشان دهنده تخمین موقعیت توسط الگوریتم EKF است و نقاط سبز رنگ نشانه های نقشه هستند. مسیرهای قرمز نشان دهنده نتایج الگوریتم RRT* است. شکل های (۱۴) تا (۱۷) نتایج شبیه سازی را نشان می دهند.



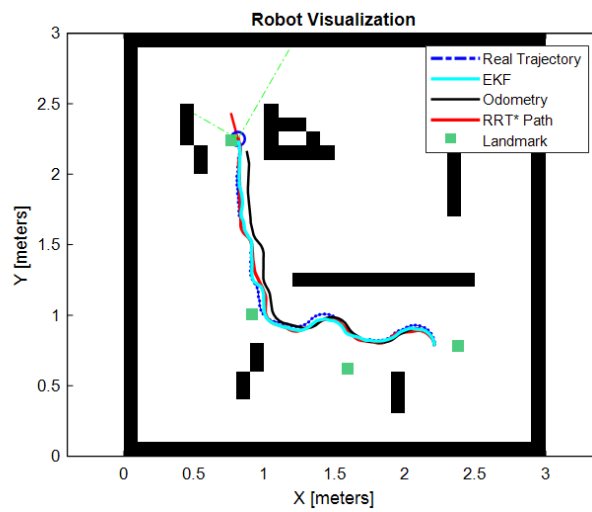
شکل ۱۴- مسیر اول تخمین زده شده توسط الگوریتم EKF.



شکل ۱۵- مسیر دوم تخمین زده شده توسط الگوریتم EKF.



شکل ۱۶- مسیر سوم تخمین زده شده توسط الگوریتم EKF.



شکل ۱۷- مسیر چهارم تخمین زده شده توسط الگوریتم EKF.

بر اساس نتایج بخش (۴)، به وضوح مشخص است که الگوریتم مسیریابی RRT^* به طور موثر مسیر بهینه را می سازد. نتایج نشان می دهد که الگوریتم RRT^* مسیر را به گونه ای طراحی می کند که الزامات ایمنی رعایت می شود و محدودیت های تعریف شده توسط محیط عملیاتی، مورد توجه ویژه واقع می شود.

۵- نتیجه گیری

در این مقاله، موقعیت یابی ربات متحرک در محیط های داخلی با استفاده از ترکیب فیلتر کالمن توسعه یافته و الگوریتم مسیریابی درخت تصادفی ارائه شده است. در روش پیشنهادی، مسئله موقعیت یابی تبدیل به یک مسئله بهینه سازی شده و سپس از RRT^* برای برنامه ریزی مسیر ربات استفاده شده است. عملکرد موقعیت یابی مبتنی بر EKF با عملکرد موقعیت یابی با ادومتری از نظر دقت و درصد خطا مورد ارزیابی قرار گرفته است. نتایج شبیه سازی ها نشان می دهند که عملکرد موقعیت یابی بر اساس الگوریتم EKF بهتر از موقعیت یابی بر اساس ادومتری است. یکی از مزایای موقعیت یابی بر اساس الگوریتم EKF در این است که این الگوریتم مانند موقعیت یابی بر اساس ادومتری تنها وابسته به اطلاعات حسگرهای درونی نیست. بنابراین خطای موقعیت یابی کاهش می یابد. سپس، روش RRT^* برای یافتن یک مسیر امکان پذیر با توجه به محیط عملیاتی معرفی می شود. روش RRT^* اهداف برنامه ریزی مسیر را برآورده می کند که این اهداف عبارتند از: (۱) مسیر بهینه در سریع ترین زمان ممکن ایجاد شود. (۲) الزامات ایمنی کاملاً رعایت شوند. (۳) محدودیت های تعریف شده توسط محیط عملیاتی، مورد توجه ویژه واقع شوند. نتایج نشان می دهند که ربات متحرک می تواند مسیر ایده آلی را در محیط های داخلی باریک پیدا کند که کارایی ناوبری مستقل و کیفیت مسیر را تضمین می کند. خطای موقعیت یابی و خطای زاویه ایجاد شده توسط روش پیشنهادی بر اساس الگوریتم EKF بسیار کمتر از الگوریتم ادومتری است و خطا با گذشت زمان افزایش نخواهد یافت. مربعات خطای موقعیت یابی $0/055$ است. در مقایسه با الگوریتم ادومتری، میانگین دقت موقعیت یابی روش پیشنهادی 73% افزایش یافته است. بهبود عملکرد از طریق شبیه سازی های متعدد تأیید می شود.

مراجع

- [1] Siegwart, R., Nourbakhsh, I.R., and Scaramuzza, D.: 'Introduction to autonomous mobile robots' MIT press, 2011.
- [2] Zheng, L., Zhan, X., Zhang, X., Wang, S., and Yuan, W.: 'Heading estimation for multimode pedestrian dead reckoning', *IEEE Sensors Journal*, vol.20, no.15, pp. 8731-8739, 2020.
- [3] Madray, I., Suire, J., Desforges, J., and Madani, M.R. 'Relative angle correction for distance estimation using K-nearest neighbors', *IEEE Sensors Journal*, vol. 20, no. 14, pp. 8155–8163, Jul. 2020
- [4] Guo, S., Zhang, Y., Gui, X., and Han, L.: 'An improved PDR/UWB integrated system for indoor navigation applications', *IEEE Sensors Journal*, vol. 20, no. 14, pp. 8046–8061, Jul. 2020.
- [5] Rauniyar, S., Bhalla, S., Choi, D., and Kim, D.: 'EKF-SLAM for Quadcopter Using Differential Flatness-Based LQR Control', *Electronics*, vol. 12, no. 5, p. 1113, 2023.
- [6] Leonard, J.J., and Durrant-Whyte, H.F.: 'Directed sonar sensing for mobile robot navigation', Springer Science & Business Media, 2012.
- [7] Chen, L., Hu, H., and McDonald-Maier, K.: 'EKF based mobile robot localization', in Proc. IEEE Int. Conf. on Industrial Technology (ICIT), Athens, Greece, 2012, pp. 149–154.

- [8] Suliman, C., Cruceru, C., and Moldoveanu, F.: 'Mobile robot position estimation using the Kalman filter', *Acta Marisiensis, Seria Technologica*, vol. 6, pp. 75–80, 2009.
- [9] Zhang, H., Chen, N., and Fan, G.: 'An Improved Localization Algorithm for Intelligent Robot', in *Proc. IEEE Int. Conf. on Artificial Intelligence and Big Data (ICAIBD)*, Chengdu, China, 2019, pp. 1–5.
- [10] Lin, H.-Y. and M.-C. Yeh, 'Drift-Free Visual SLAM for Mobile Robot Localization by Integrating UWB Technology'. *IEEE Access*, vol. 10, pp. 93636–93645, 2022.
- [11] Coelho, F.O., et al. 'EKF and computer vision for mobile robot localization', in *Proc. 13th APCA Int. Conf. Automatic Control and Soft Computing (CONTROLO)*, 2018.
- [12] Zhang, F., et al. 'A novel strategy of localization based on EKF for mobile robot', in *Proceedings of the 33rd Chinese Control Conference*. 2014.
- [13] Chen, L., H. Hu, and K. McDonald-Maier. 'EKF based mobile robot localization', in *Proc. 3rd Int. Conf. Emerging Security Technologies*, 2012.
- [14] Joon, A. and W. Kowalczyk. 'Leader Following Control of Non-holonomic Mobile Robots Using EKF-based Localization' in *2023 27th International Conference on Methods and Models in Automation and Robotics (MMAR)*. 2023.
- [15] Sun, Z., et al., 'Multi-Risk-RRT: An Efficient Motion Planning Algorithm for Robotic Autonomous Luggage Trolley Collection at Airports', *IEEE Transactions on Intelligent Vehicles*, vol. 9, pp. 3450-3463, 2024.
- [16] Mizuno, M. and T. Kubota. 'A new path planning architecture to consider motion uncertainty in natural environment', in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 2020. IEEE.
- [17] Cao, M., Zhou, X., and Ju, Y., 'Robot Motion Planning Based on Improved RRT Algorithm and RBF Neural Network Sliding', *IEEE Access*, vol. 11, pp.121295-121305, 2023.
- [18] Yu, Z., Jiang, X., and Liu, Y., 'Pose estimation of an aerial construction robot based on motion and dynamic constraints', *Robotics and Autonomous Systems*, vol. 172, p. 104591, 2024.
- [19] Brigadnov, I., Lutonin, A., and Bogdanova, K., 'Error State Extended Kalman Filter Localization for Underground Mining Environments', *Symmetry*, vol. 15, no. 2, p. 344, 2023.
- [20] Panigrahi, P.K., and Bisoy, S.K.: 'Localization strategies for autonomous mobile robots: A review', *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 8, pp. 6019-6039, 2022.
- [21] Kong, F., Chen, Y., Xie, J., Zhang, G., and Zhou, Z.: 'Mobile robot localization based on extended kalman filter', in *Proceedings of IEEE Conference*, 2006, pp. 9242–9246 .
- [22] Karaman, S. and E. Frazzoli. 'Optimal kinodynamic motion planning using incremental sampling-based methods', in *49th IEEE conference on decision and control (CDC)*. 2010.
- [23] Karaman, S. and Frazzoli, E. 'Sampling-based algorithms for optimal motion planning', *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 846–894, 2011.
- [24] Noreen, I., Khan, A., and Habib, Z., 'Optimal path planning using RRT* based approaches: a survey and future directions', *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 11, 2016.

-
- [25] Szabat, K., et al., 'A fuzzy unscented Kalman filter in the adaptive control system of a drive system with a flexible joint', *Energies*, vol. 13, no. 8, p. 2056, 2020.
- [26] Woo, R., E.-J. Yang, and D.-W. Seo, 'A fuzzy-innovation-based adaptive Kalman filter for enhanced vehicle positioning in dense urban environments', *Sensors*, vol. 19, no. 5, p. 1142, 2019
- [27] C.-Y. Yang, H. Samani, Z. Tang, and C. Li, 'Implementation of extended Kalman filter for localization of ambulance robot', *International Journal of Intelligent Robotics and Applications*, vol. 8, no. 4, pp. 960–973, Jun. 2024.
- [28] Y. Wang, Q. Zhang, and J. Liu, 'Application of a Modified IEKF Algorithm in Mobile Robot Localization', in *Proc. 2022 4th Int. Conf. on Robotics, Intelligent Control and Artificial Intelligence (RICAI)*, Guangzhou, China, Dec. 2022, pp. 53–57.



Sahand University
of Technology

DOR:

[20.1001.1.23223146.1404.12.1.4.7](https://doi.org/10.1001.1.23223146.1404.12.1.4.7)

Journal of Nonlinear
Systems in Electrical
Engineering

Vol.12, No.1

Spring and Summer 2025

ISSN: 2322 – 3146

<http://jnsee.sut.ac.ir>

Depth-Constrained Control of an Autonomous Underwater Vehicle Based on a Generalized Model Predictive Controller: Processor-in-the-Loop Implementation

Seyed Hossein Mirbagheri, Saeed Nasrollahi*

Faculty of Electrical and Computer Engineering, Malek Ashtar University of Technology,
Tehran, Iran

Abstract

Precise control of Autonomous Underwater Vehicle (AUV) is of great importance due to their wide application in various fields for tracking the path. This paper uses generalized predictive control to design and implement a processor in the loop controller for controlling the depth of an AUV. For this purpose, first, six degrees of freedom equations for the subsurface are presented and by linearizing them, the continuous-time state space equations of the AUV are extracted and by discretizing these equations, the discrete-time state space equations for the controller design are obtained. The performance of the designed controller in two unconstrained and constrained cases in the presence of actuator saturation on the linear model is investigated. Also, the performance of the proposed controller in the presence of disturbance is also investigated. For comparison, the performance of the controller designed with a linear quadratic regulator controller is also shown. Sensitivity analysis to changes in the predictive horizon and the presence of measurement noise is also shown. Finally, by implementing the processor in the loop test, the ability to implement the controller in a real-time environment has been investigated, and the results show the ability to implement the controller on hardware in real time. Keywords Autonomous Underwater Vehicle, Generalized Model Predictive Control, Processor-in-the-Loop Implementation, Depth-constrained control, Disturbance.

Keywords

Autonomous Underwater Vehicle, Generalized Model Predictive Control, Processor-in-the-Loop Implementation, Depth-constrained control, Disturbance.

1. Short Introduction

Autonomous Underwater Vehicles (AUVs) are regarded as key tools for the exploration and exploitation of underwater environments and therefore require accurate and reliable control systems. The complexity of motion dynamics, environmental disturbances, and practical constraints makes the design of advanced control strategies essential. In this context, predictive control has attracted considerable attention as an effective approach in recent studies. However, challenges such as actuator saturation, model uncertainties, and real-time implementation still remain. Consequently, evaluating the performance and practical implementability of predictive controllers is of significant importance.

2. Proposed Work

In this study, a Generalized Predictive Control (GPC) approach is developed for AUV depth control. The vehicle dynamics are modeled, linearized, and discretized to design the controller. The controller performance is evaluated under constrained and unconstrained conditions in the presence of actuator saturation, disturbances, and measurement noise, and is compared with an LQR controller. Finally, the real-time feasibility of the proposed method is assessed using a Processor-in-the-Loop test.

3. Conclusion

In this paper, a Generalized Predictive Control approach is designed and implemented for the depth control of an Autonomous Underwater Vehicle. First, the six-degree-of-freedom dynamic equations of the vehicle are developed and linearized to obtain a state-space model. The controller performance is evaluated in both constrained and unconstrained cases in the presence of actuator saturation and disturbances, and is compared with a LQR. A sensitivity analysis with respect to the prediction horizon and measurement noise is also conducted, demonstrating that increasing the prediction horizon reduces system overshoot. Finally, a Processor-in-the-Loop implementation confirms the real-time hardware feasibility of the proposed controller.



دانشگاه صنعتی شاهرود

DOR:

20.1001.1.23223146.1404.12.1.4.7

نشریه علمی-فنی در مهندسی برق

دوره ۱۲ - شماره ۱

بهار و تابستان ۱۴۰۴

صفحات ۸۱ الی ۹۹

ISSN: 2322-3146

http://jnsee.sut.ac.ir

کنترل مقید عمق یک رونده زیرسطحی مبتنی بر کنترل کننده پیش بین تعمیم یافته: پیاده سازی پردازنده در حلقه

سیدحسین میرباقری^۱، سعید نصرالهی^{۲*}

^۱ کارشناسی ارشد، مجتمع دانشگاهی برق و کامپیوتر، دانشگاه صنعتی مالک اشتر، تهران، ایران

^۲ نویسنده مسئول، استادیار، مجتمع دانشگاهی برق و کامپیوتر، دانشگاه صنعتی مالک اشتر، تهران، ایران

nasrollahi@mut.ac.ir

تاریخ پذیرش: ۱۴۰۴/۵/۶

تاریخ بازنگری: ۱۴۰۴/۱۱/۲۵ و ۱۴۰۴/۵/۵

تاریخ ارسال: ۱۴۰۳/۹/۱۸

چکیده

واژه‌های کلیدی

رونده زیرسطحی،

کنترل پیش‌بین تعمیم یافته،

آزمون پردازنده در حلقه،

کنترل مقید عمق،

اغتشاش.

کنترل دقیق زیرسطحی‌ها به دلیل کاربرد گسترده آنها در حوزه‌های مختلف به منظور ردیابی مسیر از اهمیت زیادی برخوردار است. مقاله حاضر با استفاده از کنترل پیش‌بین تعمیم یافته به طراحی و پیاده سازی پردازنده در حلقه کنترل کننده برای کنترل عمق یک رونده زیرسطحی پرداخته است. برای این منظور ابتدا معادلات شش درجه آزادی برای زیرسطحی ارائه شده و با خطی سازی آن، معادلات فضای حالت زمان پیوسته زیرسطحی استخراج شده و با گسسته سازی این معادلات، معادلات فضای حالت زمان گسسته برای طراحی کنترل کننده بدست آمده است. عملکرد کنترل کننده طراحی شده در دو حالت نامقید و مقید در حضور اشباع عملگر بر روی مدل خطی بررسی شده است. همچنین، عملکرد کنترل کننده پیشنهادی در حضور اغتشاش نیز بررسی شده است. به منظور مقایسه، عملکرد کنترل کننده طراحی شده با کنترل کننده تنظیم کننده درجه دوم خطی نیز نشان داده شده است. تحلیل حساسیت نسبت به تغییرات افق پیش‌بینی و وجود نویز اندازه گیری نیز نشان داده شده است. در انتها با پیاده سازی آزمون پردازنده در حلقه، قابلیت پیاده سازی کنترل کننده در محیط زمان واقعی بررسی شده است که نتایج، قابلیت پیاده سازی کنترل کننده بر روی سخت افزار به صورت زمان واقعی را نشان می‌دهد.

۱- مقدمه

اقیانوس‌ها حدود ۷۱٪ از سطح کره زمین را پوشانده‌اند که بیش از ۸۰٪ از حجم این اقیانوس‌ها، بدون نقشه و کاوش نشده باقی مانده است [۱]. زیرسطحی‌ها (AUV) ربات‌های زیرآبی هستند که فرآیند جمع‌آوری اطلاعات از اقیانوس‌ها را متحول کرده‌اند [۲]. اولین زیرسطحی‌ها در حوزه نظامی کاربرد داشتند ولی امروزه آنها در حوزه نفت و گاز به خصوص در عمق خیلی پایین که برای غواصان مشکل است، یا زمین شناسی و باستان شناسی زیرآب و... کاربرد دارند [۳]. به دلیل کاربردهای گسترده زیرسطحی‌ها، کنترل دقیق آن و دنبال کردن مسیر طراحی شده، از اهمیت بالایی برخوردار است به گونه‌ای که تحقیقات گسترده‌ای در این رابطه منتشر شده است.

روش‌های متفاوتی برای کنترل زیرسطحی با استفاده از کنترل کننده‌های کلاسیک و مدرن (هوشمند) ارائه شده است. در مرجع [۴]، یک کنترل کننده PID فازی تطبیقی خودتنظیم برای کنترل سمت و عمق زیرسطحی ارائه شده است. بعد از مدل‌سازی و بدست آوردن توابع تبدیل زیرسطحی، کنترل کننده ذکر شده طراحی شده و با یک PID کلاسیک مقایسه شده است. نتایج نشان‌دهنده ردیابی بهتر در ورودی‌های پله و پالس مربعی با فراجاهش کمتر است. در مرجع [۵]، یک کنترل بهینه تنظیم کننده درجه دوم (LQR) خطی برای کنترل عمق زیرسطحی طراحی شده است. ابتدا یک مدل خطی را برای کانال عمودی زیرسطحی توسعه داده شده و سپس کنترل کننده طراحی شده برای دو سیگنال مرجع بررسی شده و نشان داده شده که خروجی به خوبی دنبال می‌شود و زاویه فراز نیز به صفر میل می‌کند. در مرجع [۶] نیز با استفاده از مدل خطی مرجع [۵]، کنترل کننده طراحی شده و در حضور نویز و نامعینی، با استفاده از روش‌های بهینه‌سازی مختلف ارزیابی شده است. مرجع [۷]، یک روش نوین مبتنی بر یادگیری تقویتی برای کنترل مسیر سه بعدی وسایل نقلیه زیرآبی خودران ارائه می‌دهد. معماری پیشنهادی عامل نقاد مدل ۳ است که با ترکیب یک مدل شبکه عصبی و معماری عامل نقاد، تابع انتقال وضعیت را می‌آموزد و کنترل دقیقی را فراهم می‌کند. کنترل کننده 4ModelPPO که براین اساس طراحی شده، در مقایسه با روش‌های سنتی عملکرد بهتری در مسیرهای مختلف، به‌ویژه در حضور جریان‌های اقیانوسی، نشان می‌دهد. از نقاط قوت این مقاله می‌توان به بهبود سرعت همگرایی، افزایش دقت کنترل و استفاده از یادگیری برنامه‌ریزی شده اشاره کرد. با این حال، نقاط ضعف آن شامل عدم آزمایش عملی و برخی فرضیات محدود کننده است. در مجموع، مقاله نشان می‌دهد که هوش مصنوعی پتانسیل بالایی در حل مسائل کنترل حرکت برای وسایل نقلیه هوشمند زیرآبی دارد.

کنترل پیش‌بین نیز به‌عنوان یکی از روش‌های مؤثر در کنترل زیرسطحی‌ها مطرح شده است. در مرجع [۸]، یک کنترل کننده پیش‌بین اقتصادی به منظور بهینه‌کردن انرژی برای یک مدل زیرسطحی طراحی شده است. بهینه‌سازی انرژی با طراحی مسیر بهینه انجام شده است به گونه‌ای که مسیر حرکت زیرسطحی به دو بخش پویا و ایستا تقسیم شده است که در بخش پویا زیرسطحی باید به نقطه‌ای برسد که مسیر حرکت از آن نقطه به بعد روی خط دید طی شود که این مسیر خط دید را بخش ایستا نامیده است. در مرجع [۹]، یک کنترل کننده پیش‌بین تطبیقی بر پایه توابع لاگرانژ برای دنبال کردن مسیر در زیرسطحی طراحی شده است. کنترل کننده

¹ Autonomous underwater vehicle (AUV)

² Linear Quadratic Regulator (LQR)

³ Actor-Model-Critic

⁴ Model Proximal Policy Optimization

طراحی شده در سناریوهای مختلف با کنترل کننده پیش‌بین تطبیقی معمولی و کنترل کننده پیش‌بین معمولی مقایسه شده و کمتر بودن نوسانات و خطا را نسبت به مسیر مرجع نشان می‌دهد. در مرجع [۱۰]، یک روش کنترل پیش‌بین مدل مقاوم ۱ با حلقه‌های کنترل دوگانه ۲ برای ردیابی مسیر AUV ارائه شده است که قادر به مدیریت عدم قطعیت مدل دینامیکی و اغتشاشات تصادفی محیطی است. در این روش، یک مشاهده‌گر حالت توسعه یافته ۳ برای تخمین و جبران عدم قطعیت‌ها طراحی شده و از کنترل پیش‌بین مبتنی بر تیوب برای کاهش تأثیر اغتشاشات خارجی استفاده شده است. نتایج شبیه‌سازی‌ها نشان می‌دهند که روش پیشنهادی دقت بالاتری در کنترل مسیر سه‌بعدی AUV دارد، نوسانات مسیر را کاهش داده و در برابر اغتشاشات مقاوم‌تر از سایر روش‌ها عمل می‌کند. تحلیل پایداری مبتنی بر نظریه لیاپانوف نیز تضمین می‌کند که سیستم در شرایط نامطلوب پایدار باقی می‌ماند. باین حال، این روش هنوز در سخت‌افزار واقعی پیاده‌سازی نشده و اجرای آن نیازمند حل مسائل بهینه‌سازی پیچیده است. در مرجع [۱۱]، به طراحی کنترل کننده پیش‌بین مقید برای طراحی بخش کنترل عمق یک زیر سطحی پرداخته شده است. در این مقاله نشان داده شده است که دستور عمق داده شده که از نوع پله است به خوبی دنبال شده است؛ ولی اغتشاش و نویز بررسی نشده است.

در مقاله حاضر با استفاده از کنترل پیش‌بین تعمیم یافته (GPC4) به طراحی کنترل کننده برای کنترل عمق زیرسطحی و بررسی امکان پیاده‌سازی آن با استفاده از آزمون پردازنده در حلقه پرداخته شده است. برای این منظور ابتدا معادلات شش درجه آزادی برای یک جسم صلب دارای تقارن نوشته شده و سپس برای زیرسطحی ارائه شده است و با خطی سازی این معادلات با فرضیات مشخص، معادلات فضای حالت زمان پیوسته استخراج شده و سپس با گسسته سازی این معادلات، معادلات فضای حالت زمان گسسته برای طراحی کنترل کننده بدست آمده است. عملکرد کنترل کننده GPC در دو حالت نامقید و مقید در حضور اشباع عملگر بر روی مدل خطی بررسی شده است. به منظور مقایسه، عملکرد کنترل کننده مورد نظر با کنترل کننده تنظیم کننده درجه دوم خطی (LQR5) نیز نشان داده شده است. حساسیت سنجی کنترل کننده نسبت به پارامترهای طراحی نظیر افق پیش‌بینی و ماتریس‌های وزنی و اغتشاش با دامنه‌های مختلف و نویز اندازه‌گیری با قدرت‌های مختلف نیز انجام شده است. با آزمون پردازنده در حلقه، عملکرد کنترل کننده در محیط زمان واقعی در نرم افزار متلب تحت سناریوهای مختلف بررسی شده است.

این مقاله به پنج بخش تقسیم شده است. بخش اول به مقدمه اختصاص دارد. در بخش دوم به مدل سازی زیرسطحی پرداخته شده است. سپس در بخش سوم ابتدا ریاضیات مورد نیاز برای طراحی کنترل کننده پیش‌بین بیان شده است. در بخش چهارم به بررسی نتایج شبیه‌سازی و پیاده‌سازی آزمون پردازنده در حلقه پرداخته شده است. بخش پنجم نیز به نتیجه‌گیری و جمع بندی اختصاص دارد.

۲- مدل سازی وسیله

ساختار کلی یک زیرسطحی نوعی در شکل ۱ آمده است. معادلات شش درجه آزادی یک وسیله صلب در مرکز جرم به صورت زیر بیان می‌شود [۱۲].

¹ Robust MPC

² Dual Closed-loop Control

³ Finite-Time Extended State Observer

⁴ Generalized predictive control

⁵ Linear Quadratic Regulator

سرعت‌های زاویه‌ای حول محورهای x, y, z در دستگاه بدنی و X_G, Y_G, Z_G مختصات مرکز بویانسی نسبت به مرکز جرم در دستگاه بدنی هستند. برای طراحی کنترل کننده نیاز به خطی سازی معادلات (۲) داریم که بر اساس فرضیات مرجع [۵] خطی سازی این معادلات در کانال عمودی به صورت زیر انجام می شود.

$$\begin{aligned} \dot{\theta} &= q \\ (m - Z_{\dot{w}})\dot{w} - (mx_G + Z_{\dot{q}})\dot{q} &= Z_w U w \\ &+ (m + Z_q)Uq + Z_{\delta} U^2 \delta \\ (-M_{\dot{w}} - mx_G)\dot{w} + (I_y - M_{\dot{q}})\dot{q} &= -(z_G W - z_B)\theta \\ &+ M_w U w + (M_q - mx_G)Uq - M_{\delta} U^2 \delta \\ \dot{z} &= -U\theta + w \end{aligned} \quad (۳)$$

که فرم ماتریسی معادلات (۳) به صورت زیر بیان می شود.

$$\begin{aligned} &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & (m - Z_{\dot{w}}) & -(mx_G + Z_{\dot{q}}) & 0 \\ 0 & (-M_{\dot{w}} - mx_G) & (I_y - M_{\dot{q}}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{w} \\ \dot{q} \\ \dot{z} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & Z_w U & (m + Z_q)U & 0 \\ -(z_G W - z_B) & M_w U & (M_q - mx_G)U & 0 \\ -U & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta \\ w \\ q \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ Z_{\delta} U^2 \\ -M_{\delta} U^2 \\ 0 \end{bmatrix} \delta \end{aligned} \quad (۴)$$

برای طراحی کنترل کننده پیش بین نیاز به فرم فضای حالت رونده زیرسطحی است برای این منظور داریم:

$$\begin{aligned} \dot{x} &= A_c x + B_c u \\ y &= C_c x \end{aligned} \quad (۵)$$

که در آن

$$x = \begin{bmatrix} \theta \\ w \\ q \\ z \end{bmatrix}, u = \delta, A_c = \begin{bmatrix} 0 & 0 & 1 & 0 \\ a_{21} z_{GB} & a_{22} U & a_{23} U & 0 \\ a_{31} z_{GB} & a_{32} U & a_{33} U & 0 \\ -U & 1 & 0 & 0 \end{bmatrix}, B_c = \begin{bmatrix} 0 \\ b_1 U^2 \\ b_2 U^2 \\ 0 \end{bmatrix}, C_c = [0 \ 0 \ 0 \ 1]$$

است. رابطه بین پارامترهای رابطه (۵) و مقادیر مربوط به هندسه رونده در پیوست آمده است. مقادیر پارامترهای معادله (۵) بر اساس مرجع [۶] در جدول ۱ بیان شده است؛ که ابعاد و پارامترهای داده شده براساس زیرسطحی ارائه شده در مرجع [۱۳] است.

جدول ۱- پارامترهای مدل زیر سطحی

مقدار	ضرایب
0.0175	$a_{21}z_{GB}$
-0.2546	a_{22}
-0.7118	a_{23}
-0.052	$a_{31}z_{GB}$
0.2546	a_{32}
-0.5322	a_{33}
5	U
0.0034	b_1
0.8716	b_2

۳- کنترل پیش‌بین تعمیم یافته

در طراحی کنترل پیش‌بین تعمیم یافته نیاز به مدل فضای حالت زمان گسسته سیستم داریم، که برای این کار باید مدل فضای حالت زمان پیوسته‌ای که در بخش قبل بدست آوردیم را گسسته کنیم. برای تبدیل به فضای حالت زمان گسسته از روش اوایلر و تقریب مشتق اول در $t = kT$ استفاده می‌شود که در آن T زمان نمونه برداری است.

$$\begin{aligned} \dot{x} &\cong \frac{1}{T}(x((k+1)T) - x(kT)) \\ &= A_c x(kT) + B_c u(kT) \\ \rightarrow x((k+1)T) &= TA_c x(kT) + TB_c u(kT) + x(kT) \\ \rightarrow x((k+1)T) &= (I + TA_c)x(kT) + TB_c u(kT) \\ \rightarrow x(k+1) &= (I + TA_c)x(k) + TB_c u(k) \end{aligned} \quad (7)$$

با توجه به (۶)، رابطه بین ماتریس‌های مدل فضای حالت زمان گسسته و زمان پیوسته به صورت رابطه (۷) بیان می‌شود.

$$A_d = (I + TA_c), B_d = TB_c, C_d = C_c \quad (7)$$

مدل فضای حالت زمان گسسته سیستم به صورت (۸) بیان می‌شود.

$$\begin{aligned} x(k+1) &= A_d x(k) + B_d u(k) \\ y(k) &= C_d x(k) \end{aligned} \quad (8)$$

مدل فضای حالت سیستم به فرم رابطه (۹) بازنویسی می‌شود [۱۴]. این کار باعث اضافه شدن یک انتگرال گیر به سیستم می‌شود و به مقاوم بودن در برابر اغتشاش پله و کاهش خطای حالت ماندگار کمک می‌کند.

$$\begin{cases} \begin{bmatrix} \Delta x(k+1) \\ y(k+1) \end{bmatrix} = \begin{bmatrix} A_d & 0 \\ C_d A_d & 1 \end{bmatrix} \begin{bmatrix} \Delta x(k) \\ y(k) \end{bmatrix} + \begin{bmatrix} B_d \\ C_d B_d \end{bmatrix} \Delta u(k) \\ Y(k) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta x(k) \\ y(k) \end{bmatrix} \end{cases} \quad (9)$$

$$\rightarrow \begin{cases} X(k+1) = AX(k) + B\Delta u(k) \\ Y(k) = CX(k) \end{cases}$$

به مدل فضای حالت بیان شده در رابطه (۹)، مدل افزونه می‌گویند. براساس مدل (۹)، روابط پیش‌بینی خروجی سیستم به شرط داشتن اطلاعات کنونی در N_p گام آینده به صورت (۱۰) بیان می‌شود.

$$\bar{Y} = F\bar{X}(k) + \Phi\Delta\bar{U} \quad (10)$$

که در آن

$$\bar{Y} = \begin{bmatrix} Y(k+1) \\ Y(k+2) \\ \vdots \\ Y(k+N_p) \end{bmatrix}_{N_p \times 1}, \quad \Delta\bar{U} = \begin{bmatrix} \Delta u(k) \\ \Delta u(k+1) \\ \vdots \\ \Delta u(k+N_c) \end{bmatrix}_{N_c \times 1}, \quad F = \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^{N_p} \end{bmatrix}_{N_p \times n}$$

$$\Phi = \begin{bmatrix} CB & 0 & \dots & 0 \\ CAB & CB & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{N_p-1}B & CA^{N_p-2}B & \dots & CAB \end{bmatrix}_{N_p \times N_c}$$

n تعداد متغیرهای حالت مدل افزونه، N_p افق پیش‌بینی و N_c افق کنترل است. اساس کار کنترل پیش‌بین، مانند کنترل بهینه است با این تفاوت که در افق محدودی عمل بهینه‌سازی انجام می‌شود که این افق محدود همان افق پیش‌بینی است. افق پیش‌بین یک پارامتر مهم برای طراحی کنترل پیش‌بین است که انتخاب مناسب آن می‌تواند دقت ردیابی، سیگنال کنترل نرم‌تر و زمان واقعی بودن الگوریتم را تضمین کند [۱۵]. تابع هزینه مورد نظر به صورت (۱۱) بیان می‌شود.

$$J = \bar{E}^T Q \bar{E} + \Delta\bar{U}^T R \Delta\bar{U} \quad (11)$$

که در آن $\bar{E} = \bar{Y}_d - \bar{Y}$ و \bar{Y}_d مقادیر سیگنال مرجع در آینده، Q ماتریس جریمه خطای ردیابی و R ماتریس جریمه سیگنال کنترل است. اگر بخواهیم مقدار بهینه سیگنال کنترل را به صورتی بدست آوریم که این تابع هزینه کمینه شود کافی است، از J نسبت به \bar{U} مشتق بگیریم که در رابطه (۱۲) بدست آمده است.

$$\begin{aligned} \frac{\partial J}{\partial \Delta \bar{U}} &= \frac{(\bar{Y}_d - \bar{Y})^T Q (\bar{Y}_d - \bar{Y}) + \Delta \bar{U}^T R \Delta \bar{U}}{\partial \Delta \bar{U}} = 0 \\ &\rightarrow -2\Phi^T Q \bar{Y}_d + 2\Phi^T Q F X(k) + 2\Phi^T Q \Phi \Delta \bar{U} + 2R \Delta \bar{U} = 0 \\ &\rightarrow \Phi^T Q \Phi \Delta \bar{U} + R \Delta \bar{U} = \Phi^T Q \bar{Y}_d - \Phi^T Q F X(k) \\ &\rightarrow (\Phi^T Q \Phi + R) \Delta \bar{U} = \Phi^T Q (\bar{Y}_d - F X(k)) \\ \Delta \bar{U} &= (\Phi^T Q \Phi + R)^{-1} \Phi^T Q (\bar{Y}_d - F X(k)) \end{aligned} \quad (12)$$

در رابطه (۱۲) با حل مسئله بهینه سازی مقدار $\Delta \bar{U}$ که یک بردار از ورودی‌های آینده سیستم به طول افق کنترل است؛ بدست می‌آید که در هر لحظه اولین مقدار این بردار یعنی $\Delta U(k)$ مهم است که بر این اساس، سیگنال کنترل هر لحظه طبق رابطه (۱۳) بدست می‌آید.

$$\begin{aligned} \Delta U(k) &= U(k) - U(k-1) \\ \rightarrow U(k) &= \Delta U(k) + U(k-1) \end{aligned} \quad (13)$$

مقدار سیگنال کنترل بدست آمده در (۱۳) بدون در نظر گرفتن قید روی محاسبه شد که این موضوع در مسائل کنترل به خصوص در مسئله مورد نظر قابل قبول نیست و باید قیود مربوط به اشباع عملگر وسیله در طراحی سیگنال کنترل در نظر گرفته شود که در این صورت باید از روش‌های عددی برای حل مسئله بهینه سازی استفاده شود که این کار با استفاده از الگوریتم هیلدرث در مرجع [۱۶] انجام شده است.

در مسئله طراحی کنترل پیش‌بین، زمانی که قیدها در نظر گرفته شوند؛ مسئله به طراحی کنترل غیرخطی تبدیل می‌شود و ویژگی‌های پایداری سیستم‌های خطی تغییرناپذیر با زمان دیگر قابل اعمال نیستند. با این حال، یک خاصیت مهم کنترل پیش‌بین این است که می‌توان پایداری سیستم حلقه بسته را تحت شرایط خاصی تضمین کرد [۱۶]. کنترل پیش‌بین بر اساس اصل افق کاهنده پیاده‌سازی می‌شود به این گونه که در هر لحظه سیگنال کنترلی که با بهینه سازی تابع هزینه J در طول افق پیش‌بین بدست می‌آید یک بردار به طول افق کنترل است که درایه اول آن به سیستم اعمال می‌شود. برای اثبات پایداری تابع هزینه را به عنوان تابع لیاپانوف فرض می‌کنیم (برای راحتی کار مقادیر افق پیش‌بین و کنترل را برابر فرض می‌کنیم).

$$V_k = J_k = \sum_{k=1}^{N_p} (e_k^2 + \Delta u_k^2) \quad (14)$$

تابع بالا یک تابع مثبت معین است. بر اساس رابطه ۱۵ این تابع یک تابع کاهشی نیز است پس می‌تواند به عنوان تابع لیاپانوف سیستم برای بررسی پایداری استفاده شود.

$$\left. \begin{aligned} V_k &= \sum_{k=0}^{N_p} (e_{k+1}^2 + \Delta u_k^2) \\ V_{k+1} &= \sum_{k=1}^{N_p} (e_{k+1}^2 + \Delta u_k^2) \end{aligned} \right\} \rightarrow e_1^2 + \Delta u_0^2 + V_{k+1} = V_k \rightarrow V_{k+1} \leq V_k \quad (15)$$

اگر تابع هزینه را با افق زیاد در نظر بگیریم؛ بررسی‌ها نشان داده‌اند که رفتار کنترل کننده پیش‌بین به کنترل کننده LQR نزدیک می‌شود. از آنجایی که LQR ذاتاً پایدار است، این نتیجه را می‌توان به کنترل کننده پیش‌بین نیز تعمیم داد [۱۶]. در افق بی‌نهایت، سیستم به حالت ماندگار خود می‌رسد که در این حالت $x_k = x_{k+1}$ ، $u_k = u_{k+1}$ است و باعث صفر شدن تابع لیاپانوف می‌شود.

۴- ارزیابی نتایج و شبیه‌سازی

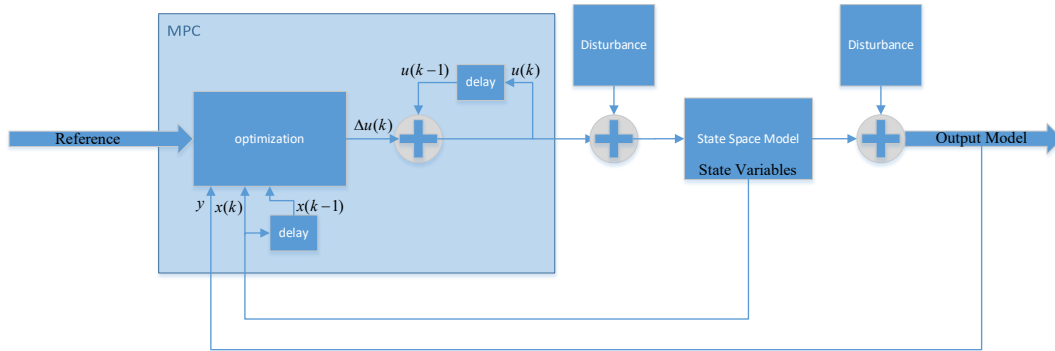
کارایی و مقاوم‌بودن کنترل کننده طراحی شده علاوه بر شبیه‌سازی، در آزمون پردازنده در حلقه تحت سناریوهای مختلف بررسی می‌شود. پارامترهای مورد استفاده در شبیه‌سازی و پردازنده در حلقه در جدول ۲ آمده است.

جدول ۲- پارامترهای طراحی

مقدار	ضرایب
100	N_p
20	N_c
0.01sec	زمان نمونه‌برداری
$100I_{N_c}$	R
I_{N_p}	Q
$\leq 30 \text{ deg}$	محدودیت اعمال ورودی

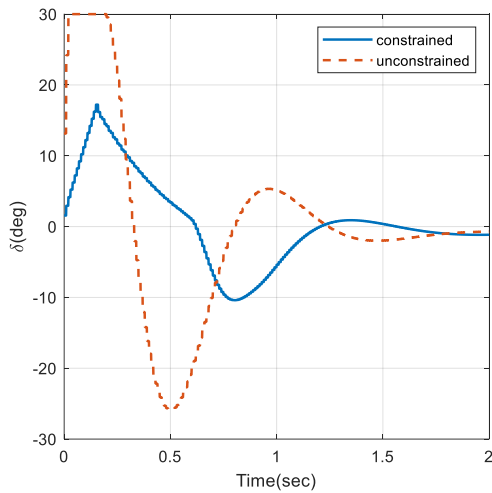
۴-۱- شبیه‌سازی

دیاگرام بلوکی شبیه‌سازی در شکل ۲ آمده است. همان‌گونه که در بخش ۳ گفته شد؛ برای طراحی کنترل کننده پیش‌بین از مدل افزونه استفاده می‌شود که متغیرهای حالت آن خروجی مورد نظر سیستم و تغییرات متغیرهای حالت مدل اصلی است. با توجه به مرجع [۱۴] کنترل کننده پیش‌بین تعمیم یافته به اغتشاش پله مقاوم است که می‌تواند در ورودی سیستم یا در خروجی سیستم لحاظ شود که این مورد در دیاگرام بلوکی شکل ۲ نیز نشان داده شده است.

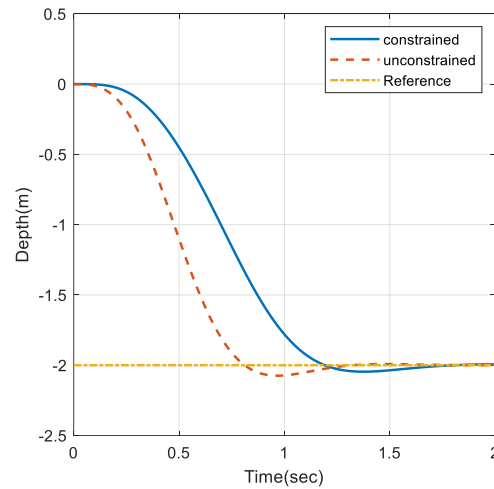


شکل ۲- دیاگرام بلوکی کنترل کننده شبیه سازی شده

با حل مسئله بهینه سازی، سیگنال کنترل بهینه بدست می‌آید و چون خروجی بهینه سازی $\Delta u(k)$ است باید برای بدست آوردن $u(k)$ ، آن را با $u(k-1)$ جمع شود. بهینه سازی می‌تواند در دو حالت مقید و نامقید صورت گیرد. در حالت مقید، قیود سیستم که، در مسئله مورد نظر همان قید اشباع عملگر سیستم هستند، در طول حل مسئله بهینه سازی که، با استفاده از روش‌های عددی حل می‌شود، لحاظ می‌گردد ولی در حالت نامقید، سیگنال کنترل بهینه از رابطه (۱۲) بدست می‌آید. خروجی سیستم به ورودی پله در دو حالت مقید و نامقید در شکل ۳ آمده است.



(ب)

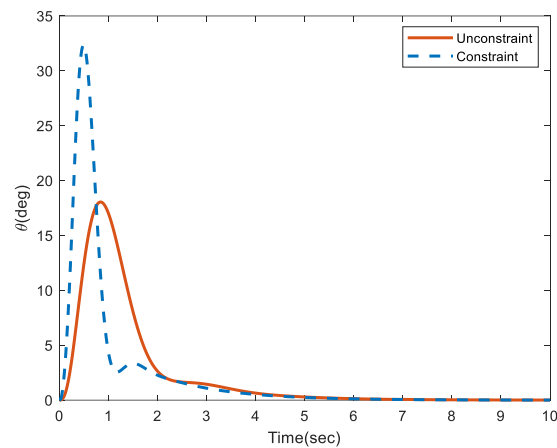


(الف)

شکل ۳- (الف) ردیابی عمق رونده زیرسطحی به ورودی پله (ب) سیگنال کنترل به ورودی پله

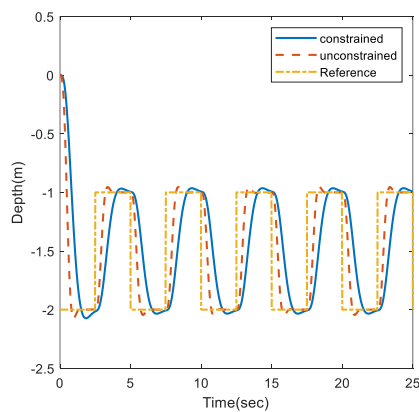
در شکل (۳) مشاهده می‌شود که در حالت نامقید سیستم، زمان نشست خروجی عمق حدوداً یک ثانیه کمتر در مقایسه با حالت مقید است. این در حالی است که سیگنال کنترل، زاویه بالک، بیشتری (در حدود ۲ برابر حالت مقید) به سیستم اعمال می‌شود و همانطور که در شکل ۳ (الف) مشاهده می‌شود در حالت نامقید سیگنال کنترل اشباع شده است. در حالت مقید با اینکه سیستم دیرتر به حالت نهایی خود می‌رسد ولی دارای سیگنال کنترل با دامنه حدوداً ۵۰ درصد حالت نامقید است و در ضمن قید محدودیت ورودی نیز

رعایت شده است. با توجه به اینکه زیرسطحی، فرمان عمق را به درستی ردیابی کرده و در عمق مشخص قرار گرفته است پس زاویه فراز نیز، که یکی از متغیرهای حالت سیستم است پایدار و در نهایت به صفر میل کرده است. در شکل ۴ پایداری زاویه فراز وسیله در دو حالت مقید و نامقید آمده است.

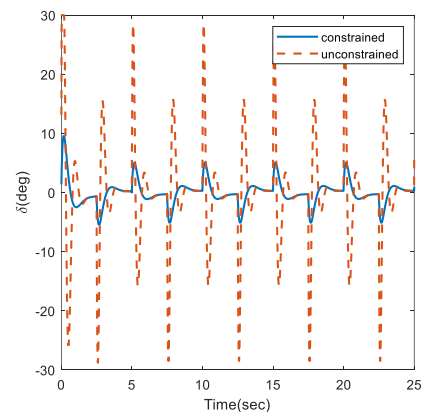


شکل ۴- زاویه فراز رونده زیرسطحی در دو حالت مقید و نامقید به ورودی پله

در ادامه برای بررسی بیشتر عملکرد کنترل کننده طراحی شده نتایج شبیه سازی برای ورودی پالس مربعی در شکل ۵ نشان داده شده است. همانطور که در شکل ۵ مشاهده می شود کنترل کننده زمانی که عمق رونده به صورت پریودیک تغییر می کند توانسته ردیابی خوبی داشته باشد.



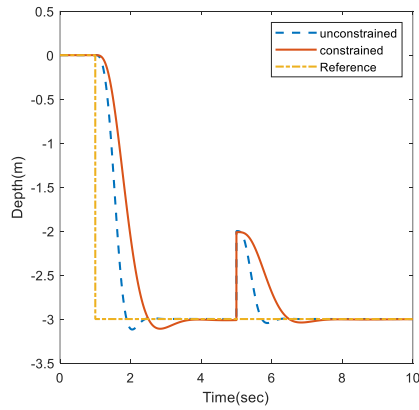
(ب)



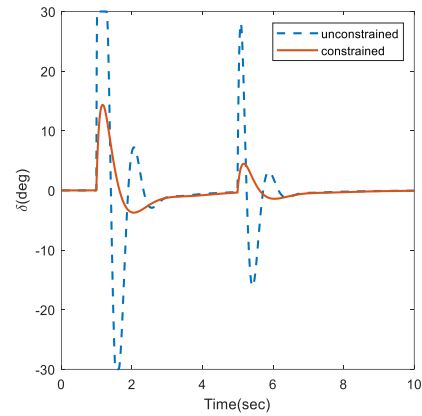
(الف)

شکل ۵- (الف) ردیابی عمق رونده زیرسطحی به ورودی پالس مربعی (ب) سیگنال کنترل به ورودی پالس مربعی

در ادامه عملکرد کنترل کننده طراحی شده نسبت به اغتشاش در شکل ۶ نشان داده شده است. در این سناریو یک اغتشاش پله دامنه واحد ثانیه ۵ به سیستم اعمال شده است. در شکل (ب) مشاهده می شود کنترل کننده توانسته اغتشاش را در کمتر از ۲ ثانیه میرا کند و ردیابی به خوبی انجام شود. در شکل ۶ (الف) مقدار زاویه بالک لازم برای میرا کردن اغتشاش نشان داده شده است.



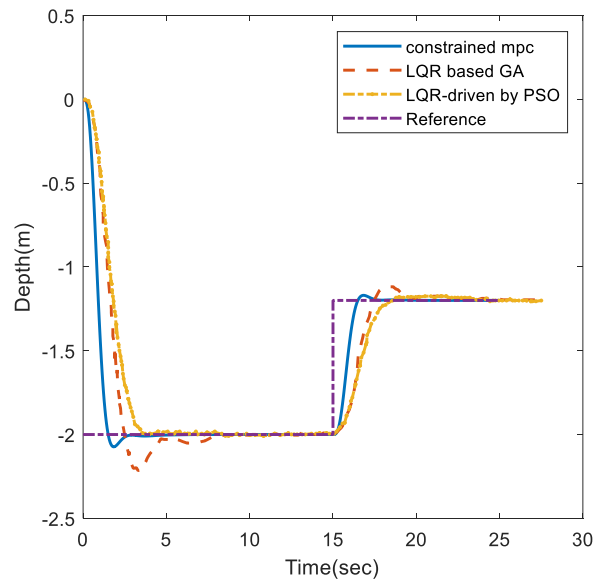
(ب)



(الف)

شکل ۶- (الف) ردیابی عمق رونده زیرسطحی در مواجهه با اغتشاش پله (ب) سیگنال کنترل در مواجهه با اغتشاش پله

در شکل ۷ عملکرد کنترل کننده طراحی شده همراه با کنترل کننده LQR مرجع [۶] نشان داده شده است. نتایج نشان می‌دهد که کنترل کننده پیش‌بین مقید در مقایسه با کنترل کننده LQR دارای زمان نشست کمتری است؛ در مورد فراجش نیز مقدار خیلی کمتری در مقایسه با LQR-GA دارد. شایان ذکر است که ضرایب کنترل کننده LQR با الگوریتم ژنتیک و الگوریتم ازدحام ذرات بهینه شده است.

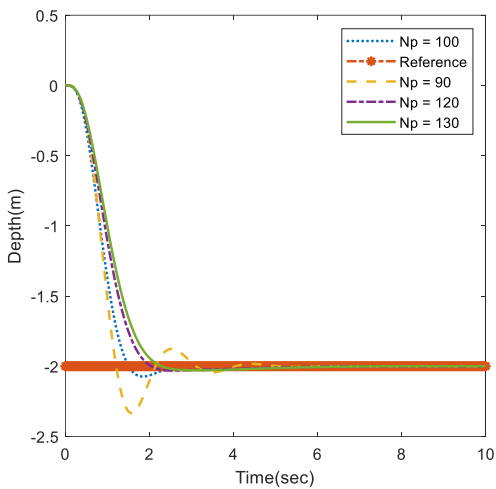


شکل ۷- مقایسه کنترل پیش‌بین و کنترل کننده LQR در کنترل عمق زیرسطحی

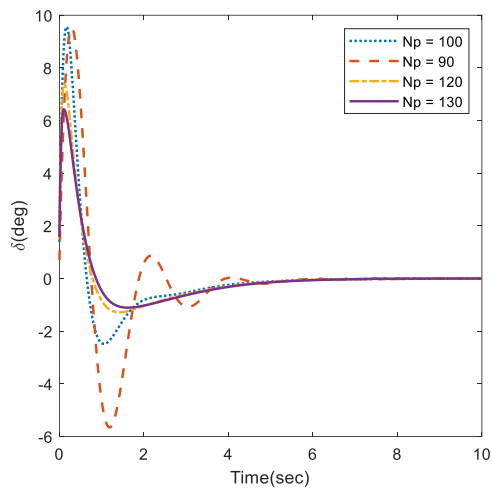
۴-۲- تحلیل حساسیت

مقاوم بودن کنترل کننده طراحی شده نسبت به تغییرات پارامترهای طراحی و در شرایط محیطی مختلف با نویز اندازه‌گیری و یا اغتشاش با دامنه‌های مختلف باید قبل از به کارگیری آن بررسی شود تا از خسارات جلوگیری شود. برای این منظور در شکل ۸

تغییرات افق پیش‌بینی در ردیابی خروجی، در شکل ۹ تغییرات افق کنترل در ردیابی خروجی و در شکل ۱۰ وجود نویز سفید اندازه‌گیری با قدرت‌های ۰/۰۵ و ۰/۰۱ بررسی شده است.

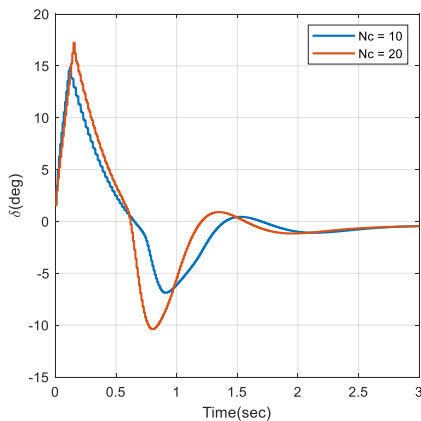


(ب)

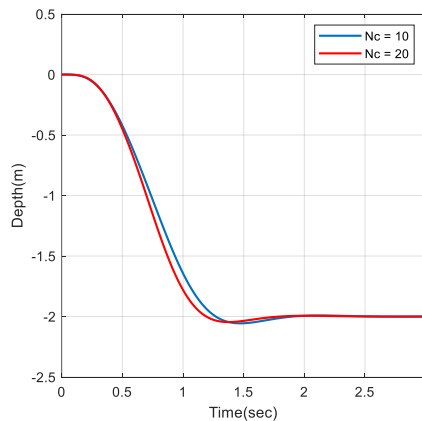


(الف)

شکل ۸- (الف) ردیابی عمق با تغییرات افق پیش‌بینی (ب) سیگنال کنترل با تغییرات افق پیش‌بینی

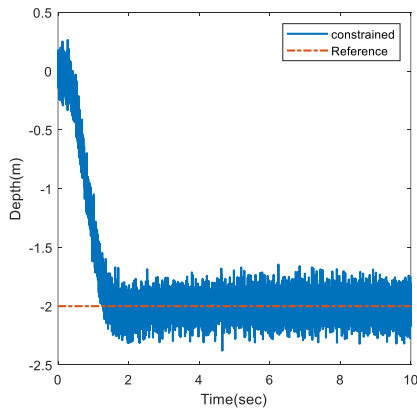


(ب)

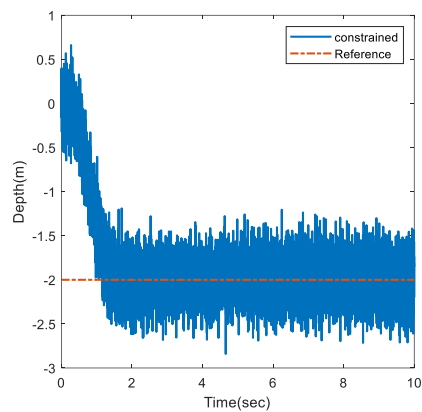


(الف)

شکل ۹- (الف) ردیابی عمق با تغییرات افق کنترل (ب) سیگنال کنترل با تغییرات افق کنترل



(ب)



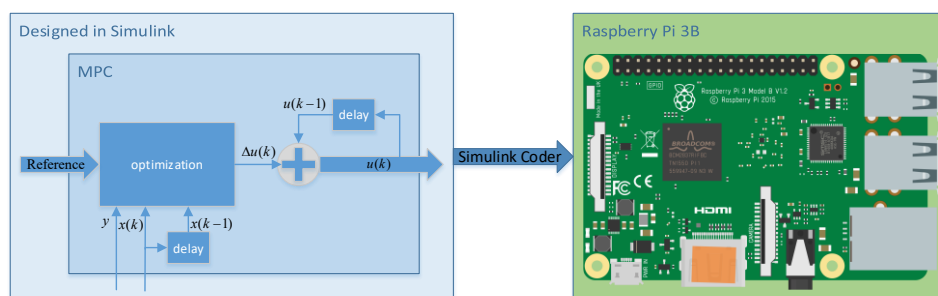
(الف)

شکل ۱۰- (الف) ردیابی عمق با وجود نویز سفید با قدرت 0.01 (ب) ردیابی عمق با وجود نویز سفید با قدرت 0.05

مشاهده می‌شود که در شکل ۸ (الف) با افزایش افق پیش‌بینی تغییرات فراجهدش سیستم کم شده است و چون با این افزایش، زمان بیشتری از آینده سیستم در بهینه‌سازی لحاظ می‌شود که مطابق شکل ۸ (ب) باعث کاهش سیگنال کنترل نیز شده است. در شکل ۹ مشاهده می‌شود که با کاهش افق کنترل از ۲۰ به ۱۰ خروجی سیستم کندتر و دارای زمان نشست بیشتری می‌شود ولی در سیگنال کنترل، دامنه کمتری اعمال شده است. در شکل ۱۰ وجود نویز اندازه‌گیری نشان داده شده است که در حضور نویز با قدرت ۰/۰۱ ردیابی با دقت ۰/۵ متر و در حضور نویز با قدرت ۰/۰۵ ردیابی با دقت حدود ۱ متر خطا انجام می‌شود.

۳-۴- پردازنده در حلقه

برای پیاده‌سازی پردازنده در حلقه از برد رزبری پای 3B و کتابخانه رزبری پای برای سیمولینک^۱ استفاده شده است. به این صورت که الگوریتم کنترلی در محیط سیمولینک طراحی شده و توسط ابزار کدکننده سیمولینک^۲ برای برد رزبری پای کامپایل و در آن اجرا می‌شود که مطابق مرجع [۱۷] انجام می‌شود. دیاگرام بلوکی این فرآیند در شکل ۱۱ نشان داده شده است.



شکل ۱۱- دیاگرام بلوکی اجرای الگوریتم کنترلی در پردازنده هدف

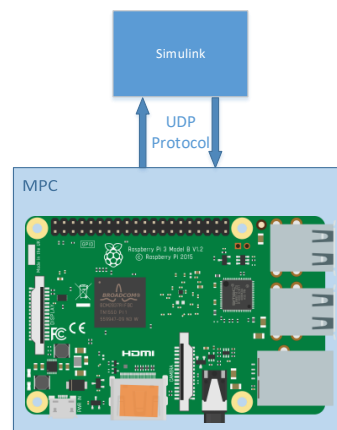
¹ Simulink

² Simulink Coder

پس از این فرآیند باید در محیط سیمولینک اطلاعات لازم برای کنترل کننده آماده شود و برای آن ارسال گردد و بعد از انجام محاسبات در کنترل کننده، سیگنال کنترل کننده در سیمولینک دریافت گردد. برای ارسال و دریافت اطلاعات بین سیمولینک و برد رزبری پای از پروتکل UDP استفاده شده است که دیاگرام بلوکی این فرآیند در شکل ۱۲ (الف) و تصویر تجهیزات در شکل ۱۲ (ب) آمده است.



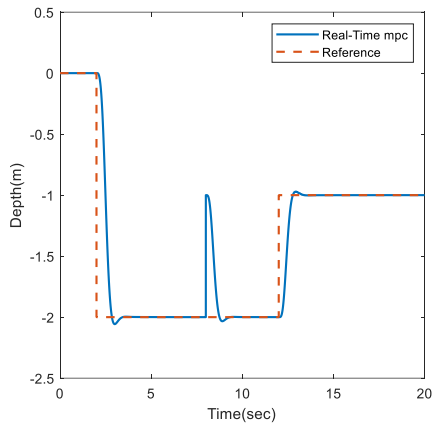
(ب)



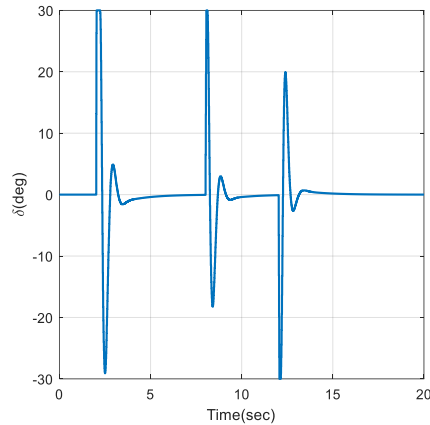
(الف)

شکل ۱۲- (الف) دیاگرام بلوکی اجرای زمان حقیقی کنترل پیش بین (ب) اجرای زمان حقیقی کنترل پیش بین

پیاده سازی پردازنده در حلقه برای دو ورودی پله و پالس مربعی انجام شده است. خروجی پردازنده در حلقه به ورودی پله همراه با اغشاش در شکل ۱۳ و به ورودی پالس مربعی در شکل ۱۴ آمده است. نتایج نشان می دهد که الگوریتم کنترل پیش بین پیاده سازی شده در برد رزبری پای سیگنال های مرجع را به خوبی دنبال کرده و نیز به اغشاش پله مقاوم است.

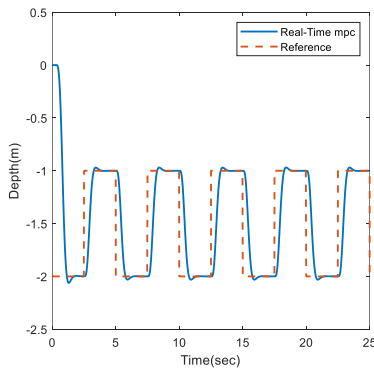


(ب)

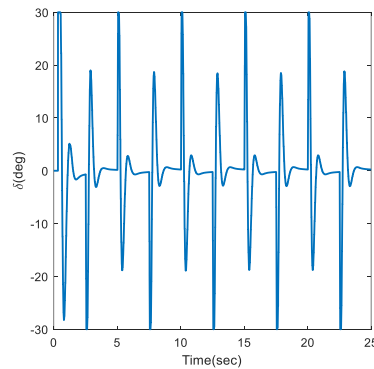


(الف)

شکل ۱۳- پیاده سازی پردازنده در حلقه: (الف) ردیابی عمق رونده زیرسطحی به ورودی پله در مواجه با اغتشاش به صورت زمان حقیقی (ب) سیگنال کنترل به ورودی پله در مواجه با اغتشاش به صورت زمان حقیقی



(ب)



(الف)

شکل ۱۴- پیاده سازی پردازنده در حلقه: (الف) ردیابی عمق رونده زیرسطحی به ورودی پالس مربعی به صورت زمان حقیقی (ب) سیگنال کنترل به ورودی پالس مربعی به صورت زمان حقیقی

۵- نتیجه گیری و جمع بندی

در این مقاله با استفاده از کنترل پیش‌بین تعمیم یافته به طراحی و پیاده سازی سخت افزاری سیستم برای کنترل عمق رونده زیرسطحی پرداخته شد. برای این منظور ابتدا معادلات شش درجه آزادی برای زیرسطحی توسعه پیدا کرد و با خطی سازی این معادلات، معادلات فضای حالت زمان پیوسته زیرسطحی استخراج شد. عملکرد کنترل کننده طراحی شده در دو حالت نامقید و مقید در حضور اشباع عملگر بر روی مدل خطی بررسی شد. همچنین، عملکرد کنترل کننده پیشنهادی در حضور اغتشاش نیز بررسی شد. به منظور

مقایسه، عملکرد کنترل کننده طراحی شده همراه با کنترل کننده تنظیم کننده درجه دوم خطی نشان داده شد. یک آنالیز حساسیت نسبت به تغییرات افق پیش بینی و وجود نویز اندازه گیری نیز انجام شد. نتایج نشان می دهد با افزایش افق پیش بینی تغییرات فراجهبش سیستم کم می شود. در انتها با پیاده سازی آزمون پردازنده در حلقه قابلیت پیاده سازی کنترل کننده در محیط زمان واقعی بررسی شد. نتایج نشان می دهد که کنترل کننده مقید ارائه شده قابلیت پیاده سازی بر روی سخت افزار به صورت زمان واقعی را دارد.

پیوست

رابطه بین مقادیر پارامترهای رابطه (۵) و مقادیر مربوط به هندسه رونده زیرسطحی:

$$Dv = (m - Z_{\dot{w}})(I_y - M_{\dot{q}}) - (mx_G + Z_{\dot{q}})(mx_G + M_{\dot{w}})$$

$$z_{GB} = z_G - z_B n$$

$$a_{21} = -\frac{(mx_G + Z_{\dot{q}})W}{Dv}$$

$$a_{22} = \frac{(I_y - M_{\dot{q}})Z_w + (mx_G + Z_{\dot{q}})M_w}{Dv}$$

$$a_{23} = \frac{(I_y - M_{\dot{q}})(m + Z_q) + (mx_G + Z_{\dot{q}})(M_q - mx_G)}{Dv}$$

$$a_{31} = \frac{-(mx_G - Z_{\dot{q}})W}{Dv}$$

$$a_{32} = \frac{(m - Z_{\dot{w}})M_w + (mx_G + M_{\dot{w}})M_{\delta}}{Dv}$$

$$a_{33} = \frac{(m - Z_{\dot{w}})(M_q - mx_G) + (mx_G + M_{\dot{w}})(m - Z_q)}{Dv}$$

$$b_1 = \frac{(I_y - M_{\dot{q}})Z_{\delta} + (mx_G + Z_{\dot{q}})M_{\delta}}{Dv}$$

$$b_2 = \frac{(I_y - Z_{\dot{w}})M_{\delta} + (mx_G + M_{\dot{q}})Z_{\delta}}{Dv}$$

منابع

- [1] X. Y. Jing Yan, Haiyan Zhao, Xiaoyuan Luo, Xinping Guan, *Autonomous Underwater Vehicles (Cognitive Intelligence and Robotics)*, Springer, 2021.
- [2] N. Cruz, *Autonomous Underwater Vehicles*, IntechOpen, 2011.
- [3] F. Fanelli, *Development and Testing of Navigation Algorithms for Autonomous Underwater Vehicles (Springer Theses)*, Springer, 2020.
- [4] M. H. Khodayari and S. Balochian, "Modeling and control of autonomous underwater vehicle (AUV) in heading and depth attitude via self-adaptive fuzzy PID controller," *Journal of Marine Science and Technology*, vol. 20, no. 3, pp. 559-578, 2015.
- [5] N. Syahroni, Y. B. Seo, and J. W. Choi, "Depth Control of Autonomous Underwater Vehicle Based on Open Control Platform," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 3707-3712, 2008.

- [6] M. T. Muhssin ,M. N. Ajaweed, and S. K. Khalaf, "Optimal control of underwater vehicle using LQR controller driven by new matrix decision control algorithm," *International Journal of Dynamics and Control*, vol. 11, no. 6, pp. 2911-2923, 2023.
- [7] Ma, D., Chen, X., Ma, W., Zheng, H., & Qu, F. "Neural Network Model-Based Reinforcement Learning Control for AUV 3-D Path Following, " *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 893-904, 2024.
- [8] N. Yang, D. Chang, M .Johnson-Roberson, and J. Sun, "Energy-Optimal Control for Autonomous Underwater Vehicles Using Economic Model Predictive Control," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 6, pp. 2377-2390, 2022.
- [9] W. Wang, J. Yan, H. Wang, H. Ge, Z .Zhu, and G. Yang, "Adaptive MPC trajectory tracking for AUV based on Laguerre function," *Ocean Engineering*, vol. 261, p.111870, 2022.
- [10] Z. Yan, J. Yan, S. Cai, Y. Yu, and Y. Wu, "Robust MPC-based trajectory tracking of autonomous underwater vehicles with model uncertainty," *Ocean Engineering*, vol. 286, p. 115617, 2023.
- [11] S. Li, S. Wang, and X. Luo, "Depth control of autonomous underwater vehicles based on constrained model predictive control, " *Proc. 42nd Chinese Control Conference (CCC)*, Tianjin, China, Jul. 2023 .
- [۱۲] س. ازگلی، م. ر. عاروان، مدل‌سازی و شبیه‌سازی سامانه‌های متحرک، انتشارات یامهدی، ۱۳۸۹.
- [13] A. J. Healey, P. A. Papoulias, and R. Cristi, "Design and Experimental Verification of a Model Based Compensator for Rapid Auv Depth Control," presented at the Proceedings of the 6th International Symposium on Unmanned Untethered Submersible Technology, 1989 .
- [14] J. A. Rossiter, *A First Course in Predictive Control*, (Second Edition), CRC Press, 2018.
- [15] H. Wang, Q. Wang ,W. Chen, L. Zhao, and D. Tan, "Path tracking based on model predictive control with variable predictive horizon," *Transactions of the Institute of Measurement and Control*, vol. 43, no. 12, pp. 2676-2688, 2021.
- [16] L. Wang, *Model Predictive Control System Design and Implementation Using MATLAB®* (Advances in Industrial Control), Springer, 2009.
- [17] Raspberry pi support from Simulink. Available: <https://www.mathworks.com/hardware-support/raspberry-pi-simulink.html>

Vehicle Detection from Images Using a Modified YOLOv4 Algorithm with High Accuracy in Deep Learning



Sahand University
of Technology

DOR:

[20.1001.1.23223146.1404.12.1.5.8](https://doi.org/10.1001.1.23223146.1404.12.1.5.8)

Journal of Nonlinear
Systems in Electrical
Engineering

Vol.12, No.1

Spring and Summer 2025

ISSN: 2322 – 3146

<http://jnsee.sut.ac.ir>

Javad Sayyadi, Mahdi nangir *, Behzad Mozaffari Tazekand, Mahmood Mohassel Feghh, Hamid sayyadi

Faculty of Electrical & Computer Engineering, University of Tabriz, Tabriz, Iran
Emails: J.sayyadi1400@ms.tabrizu.ac.ir -Nangir@tabrizu.ac.ir- Mozaffari@tabrizu.ac.ir-
Mohasselfegghi@tabrizu.ac.ir -H.sayyadi1400@ms.tabrizu.ac.ir

Abstract

Vehicle detection from images is one of the key applications of computer vision in fields such as military systems, urban surveillance, transportation, and safety. In this study, an improved version of the YOLOv4 algorithm is proposed for vehicle detection. The main innovation lies in the combination of two optimizers, ADAM and SGDM, to enhance the training speed and stability, along with the design of a lightweight post-processing algorithm, called IASD, for eliminating redundant and noisy bounding boxes. The proposed method was evaluated on a large-scale dataset, and the results showed that, compared with the original YOLOv4, significant improvements were achieved in both detection accuracy and speed. This outcome indicates that the proposed version can be a suitable option for real-time and resource-constrained systems such as UAVs and surveillance platforms. Moreover, the introduction of a lightweight and hardware-friendly post-processing algorithm distinguishes this research from similar studies. The results of this work can serve as a foundation for developing safer and more efficient applications in intelligent transportation and autonomous systems.

Keywords

Vehicle detection, YOLOv4, ADAM, SGDM, IASD algorithm

1. Introduction

Vehicle detection is a critical task in computer vision for applications like surveillance and autonomous systems. While YOLOv4 offers a balance of speed and accuracy, newer versions demand significant computational resources. This paper aims to enhance YOLOv4's performance for resource-constrained, real-time systems by addressing its training stability and post-processing efficiency. The proposed modifications focus on optimizing the learning process and refining detection outputs to achieve higher accuracy without heavy hardware reliance.

2. Proposed Work and Methodology (including comprision, simulation/experimental results and discussion)

The core of the proposed work involves two key innovations integrated into the YOLOv4 framework. First, a novel hybrid optimizer combining ADAM and SGDM is employed to accelerate training convergence and improve stability, as evidenced by a faster and lower training loss curve compared to the standard model. Second, a lightweight post-processing algorithm named IASD is designed to filter out redundant, noisy, and low-confidence bounding boxes through steps like confidence thresholding, Non-Maximum Suppression (NMS), and box merging. This significantly enhances detection precision. The method was evaluated on a large-scale dataset. Results demonstrate that the proposed YOLOv4+IASD model achieves a notable improvement over the original YOLOv4, increasing mAP (IoU=0.5) from 83.3% to 90.1% and mAP (IoU=0.5:0.95) from 66.0% to 68.4%, while maintaining a practical speed of 70 FPS on CPU-limited hardware. This performance is competitive with newer models like YOLOv5/v7/v8, which require GPUs for their superior FPS. The discussion highlights that this makes the proposed model a suitable and efficient choice for embedded and real-time applications such as UAVs and urban surveillance systems.

3. Conclusion

This paper presented an enhanced YOLOv4 algorithm for vehicle detection, achieving a significant balance between accuracy and computational efficiency. The integration of a hybrid ADAM-SGDM optimizer and the novel IASD post-processing algorithm led to measurable improvements in both mAP and training convergence speed. While newer YOLO variants excel on GPU hardware, the proposed method's lower resource requirements make it exceptionally viable for real-time, embedded applications like drones and intelligent transportation systems, bridging the gap between high performance and practical deployability.



دانشگاه صنعتی تبریز

DOR:

20.1001.1.23223146.1404.12.1.5.8

نشریه علمی-فنی مهندسی برق

دوره ۱۲ - شماره ۱

بهار و تابستان ۱۴۰۴

صفحات ۱۰۰ الی ۱۱۷

ISSN: 2322-3146

http://jnsee.sut.ac.ir

تشخیص وسیله نقلیه از تصاویر با بهره‌گیری از الگوریتم تغییر یافته YOLOv4 با دقت بالا در یادگیری عمیق

جواد صیادی^۱، مهدی نان‌گیر^۲، بهزاد مظفری تازه‌کند^۳، محمود محصل‌فقهی^۴، حمید صیادی^۵

۱- فارغ التحصیل کارشناسی ارشد، دانشکده مهندسی برق و کامپیوتر - دانشگاه تبریز - ایران

j.sayyadi1400@ms.tabrizu.ac.ir

۲- دانشیار، دانشکده مهندسی برق و کامپیوتر - دانشگاه تبریز - ایران

nangir@tabrizu.ac.ir

۳- استاد، دانشکده مهندسی برق و کامپیوتر - دانشگاه تبریز - ایران

Mozaffary@tabrizu.ac.ir

۴- دانشیار، دانشکده مهندسی برق و کامپیوتر - دانشگاه تبریز - ایران

mohasselfeghi@tabrizu.ac.ir

۵- فارغ التحصیل کارشناسی ارشد، دانشکده مهندسی برق و کامپیوتر - دانشگاه تبریز - ایران

h.sayyadi1400@ms.tabrizu.ac.ir

تاریخ پذیرش: ۱۴۰۴/۲/۱۳

تاریخ بازنگری: ۱۴۰۴/۲/۴

تاریخ ارسال: ۱۴۰۴/۴/۲۷

چکیده

تشخیص وسایل نقلیه از تصاویر یکی از کاربردهای مهم بینایی ماشین در حوزه‌هایی چون نظامی، نظارت شهری، حمل‌ونقل و ایمنی است. در این پژوهش یک نسخه بهبودیافته از الگوریتم YOLOv4 برای تشخیص وسایل نقلیه ارائه می‌شود. نوآوری اصلی در ترکیب دو بهینه‌ساز ADAM و SGDM برای افزایش سرعت و پایداری آموزش و طراحی یک الگوریتم پس‌پردازش سبک‌وزن با عنوان IASD برای حذف جعبه‌های تکراری و نویزی است. روش پیشنهادی بر روی یک مجموعه داده بزرگ ارزیابی شد و نتایج نشان داد که در مقایسه با YOLOv4 اصلی، هم از نظر دقت و هم از نظر سرعت تشخیص، بهبود قابل توجهی حاصل شده است. این دستاورد نشان می‌دهد که نسخه پیشنهادی می‌تواند گزینه‌ای مناسب برای سامانه‌های بلادرنگ و کم‌منبع مانند پهپادها و سیستم‌های نظارتی باشد. از سوی دیگر، ارائه یک الگوریتم پس‌پردازش سبک و قابل پیاده‌سازی در سخت‌افزارهای محدود، این پژوهش را از کارهای مشابه متمایز می‌سازد. نتایج این مطالعه می‌تواند به‌عنوان مبنایی برای توسعه کاربردهای ایمن‌تر و کارآمدتر در حوزه حمل‌ونقل هوشمند و سامانه‌های خودکار مورد استفاده قرار گیرد.

واژه‌های کلیدی

تشخیص وسیله نقلیه،

YOLOv4،

ADAM،

SGDM،

الگوریتم IASD

۱- مقدمه

در دهه اخیر، با پیشرفت فناوری و توسعه روش‌های یادگیری عمیق، تشخیص اشیا در تصاویر به یکی از مسائل مهم در حوزه هوش مصنوعی تبدیل شده است. تشخیص وسایل نقلیه از تصاویر، به خصوص از منابع هوایی که برای مانیتورینگ و نظارت در محیط‌های با دسترسی محدود استفاده می‌شوند، یکی از موضوعات مهم در این حوزه محسوب می‌شود [۱]. در این مقاله، به معرفی دو روش برای تشخیص وسایل نقلیه از تصاویر با استفاده از الگوریتم‌های تغییر یافته YOLOv4 در یادگیری عمیق می‌پردازیم. YOLOv4 یکی از روش‌های معروف و قدرتمند در زمینه تشخیص اشیا است که با استفاده از شبکه‌های عصبی پیچشی، قادر است اشیا مختلف را با دقت و سرعت بالا تشخیص دهد [۲].

در این پژوهش، الگوریتم YOLOv4 با هدف تشخیص وسایل نقلیه در تصاویر بهبود یافته و نسخه‌ای ارتقاء یافته از آن ارائه شده است. این الگوریتم تغییر یافته، با بهره‌گیری از تکنیک‌های پیشرفته‌تر در آموزش و بهینه‌سازی شبکه‌های عصبی، دقت و سرعت بیشتری را در تشخیص وسایل نقلیه ارائه می‌دهد. با وجود معرفی نسخه‌های جدید YOLO مانند YOLOv7, YOLOv8 و اخیراً YOLOv13، انتخاب YOLOv4 به دلیل تعادل میان دقت، سرعت و نیازمندی سخت‌افزاری در سامانه‌های تعبیه شده همچنان رایج است.

در این مقاله، ابتدا مسئله تشخیص وسیله نقلیه از تصاویر معرفی می‌شود. سپس الگوریتم تغییر یافته YOLOv4 توضیح داده شده و آزمایش‌ها و نتایج حاصل ارائه می‌گردد. در پایان نیز نتایج مقایسه و جمع‌بندی می‌شوند. این مقاله می‌تواند به عنوان یک مبنای قوی برای تحقیقات آتی در زمینه تشخیص وسایل نقلیه از تصاویر با استفاده از یادگیری عمیق مورد استفاده قرار گیرد.

در این الگوریتم تغییر یافته، دو نوآوری ارائه شده است. اولین نوآوری مربوط به تابع پس‌پردازش است که به YOLOv4 اضافه شده است. این تابع پس از تشخیص اولیه وسایل نقلیه، اقدام به اصلاح و بهبود نتایج می‌کند. این تابع شامل مراحل حذف تشابه‌های متوالی، ادغام جعبه‌های مجاور و حذف برجسب‌های تکراری است. این تابع باعث افزایش دقت و سرعت تشخیص می‌شود و نتایج بهتری را ارائه می‌دهد [۳].

دومین نوآوری ارائه شده در این مقاله مربوط به تنظیمات الگوریتم‌های ADAM و SGDM است که سبب افزایش دقت و سرعت در YOLOv4 می‌شود. الگوریتم ADAM یک الگوریتم بهینه‌سازی است که برای بهبود سرعت یادگیری شبکه‌های عصبی مورد استفاده قرار می‌گیرد [۴]. همچنین، الگوریتم SGDM یک ترکیب از الگوریتم‌های استوانه‌ای و نزولی‌گرادیان است که باعث رشد چشمگیر سرعت و دقت الگوریتم می‌شود [۵].

پژوهش حاضر، نتایج آزمایش‌های خود را بر روی یک مجموعه داده بزرگ از تصاویر نشان می‌دهد. نتایج بیانگر آن است که الگوریتم تغییر یافته باعث بهبود قابل ملاحظه دقت و سرعت تشخیص وسیله نقلیه می‌گردد. با توجه به این نکات، در این مقاله به بررسی و ارائه الگوریتم تغییر یافته‌ای برای تشخیص وسایل نقلیه از تصاویر می‌پردازیم. این الگوریتم با استفاده از تقسیم تصویر به مناطق کوچک و اختصاص برجسب به هر منطقه، قادر است نوع و موقعیت وسیله نقلیه را با دقت بالا شناسایی نماید. مزایای استفاده از این الگوریتم شامل دقت و سرعت بالا، قابلیت تشخیص چند وسیله نقلیه هم‌زمان و توانایی استفاده در برنامه‌های وسیعی از جمله نظامی، حمل‌ونقل و نظارت شهری است. الگوریتم تغییر یافته در این پژوهش توانسته است نتایج قابل قبولی را در تشخیص وسایل نقلیه از تصاویر داشته باشد و همچنین می‌تواند به عنوان یک راه‌حل کارآمد در بسیاری از برنامه‌های کاربردی استفاده شود [۶].

با وجود معرفی نسخه‌های جدید YOLO مانند YOLOv7 و YOLOv8، انتخاب YOLOv4 همچنان در بسیاری از سامانه‌های تعبیه‌شده و کاربردهای صنعتی رایج است، زیرا تعادلی مناسب میان سرعت، دقت و نیازمندی سخت‌افزاری برقرار می‌سازد. نوآوری اصلی این پژوهش در دو بخش خلاصه می‌شود: نخست، ترکیب دو الگوریتم بهینه‌ساز ADAM و SGDM به گونه‌ای که مزایای هر دو در آموزش شبکه حفظ شود و هم‌زمان دقت و سرعت افزایش یابد. دوم، معرفی یک تابع پس‌پردازش سبک‌وزن با عنوان IASD که قادر است جعبه‌های محدودکننده تکراری و نویزی را حذف کرده و دقت تشخیص را بهبود دهد. این دو رویکرد، الگوریتم YOLOv4 را برای شرایط عملکرد بلادرننگ در سامانه‌های کم‌منبع بهینه می‌سازد و آن را از نسخه‌های استاندارد متمایز می‌کند.

۲- پیشینه پژوهش

مقاله [۶] به بهبود دقت شناسایی اشیاء در شبکه‌های عصبی کانولوشنی می‌پردازد. هدف اصلی طراحی این مقاله، ایجاد یک سرعت عملیاتی بالا، برای یک تشخیص دهنده اشیاء در سیستم‌های تولید و بهینه‌سازی برای محاسبات موازی است که با شاخص نظری حجم کم محاسبات در تضاد است. هدف این مطالعه همچنین بررسی تأثیر روش‌های جدید Bag-of- و Bag-of-Freebies و Specials در طول آموزش شناسایی اشیاء توسط تشخیص دهنده است.

مقاله [۷] یک الگوریتم تشخیص ماسک صورت و تشخیص استاندارد پوشیدن ماسک بر اساس YOLO-v4 بهبود یافته را پیشنهاد می‌دهد. نویسندگان یک CSPDarkNet53¹ بهبود یافته را در شبکه عصبی به عنوان ورودی اصلی معرفی می‌کنند که هزینه محاسباتی را کاهش می‌دهد و توانایی یادگیری را بهبود می‌بخشد. آن‌ها همچنین الگوریتم تغییر اندازه تصویر را قابل تطبیق معرفی می‌کنند تا به طور مؤثر محاسبات تکراری را کاهش دهد.

YOLOv4-tiny یک روش تشخیص سریع اشیاء بر پایه YOLOv4-tiny است که ساختار شبکه را ساده‌تر می‌کند و پارامترهای محاسباتی را کاهش می‌دهد. این روش از دو ماژول ResBlock-D به جای دو ماژول CSPBlock در YOLOv4-tiny² استفاده می‌کند که پیچیدگی محاسباتی را کاهش می‌دهد. همچنین یک بلوک کمکی طراحی شده است تا اطلاعات بیشتری از اشیاء استخراج کند و خطاهای تشخیص را کاهش دهد [۸].

در پژوهش حاضر [۹] روش تشخیص کلاه ایمنی بر اساس YOLOv4 بهبود یافته بررسی می‌شود. این روش با استفاده از الگوریتم K-means³، داده‌ها را خوشه‌بندی کرده و اطلاعات لبه‌ای هدف را به دست می‌آورد. در فرایند آموزش شبکه از استراتژی آموزش چند مقیاسی استفاده می‌شود تا قابلیت سازگاری مدل با مقیاس‌های مختلف تشخیص، افزایش یابد. نتایج آزمایش‌ها نشان می‌دهند که مقدار MAP⁴ مدل به ۹۲.۸۹٪ می‌رسد، سرعت تشخیص به ۱۵ فریم بر ثانیه می‌رسد که دقت و سرعت تشخیص آن نسبت به YOLOv4 بهبود یافته‌اند.

¹ Cross-Stage Partial DarkNet 53

² You Only Look Once version 4-tiny

³ Kluster Means

⁴ mean Average Precision

در این مقاله [۱۰] یک روش تشخیص برای شکاف‌های پل بر اساس YOLOv4-FPM با ماژول Feature Pyramid پیشنهاد می‌شود تا به طور مؤثر شکاف‌ها در سازه‌های پل شناسایی شوند. YOLOv4-FPM از چارچوب تشخیص اشیا YOLOv4 به همراه ماژول Feature Pyramid برای افزایش عملکرد تشخیص شکاف‌های کوچک و بزرگ استفاده می‌کند.

این روش از مجموعه داده بزرگ از تصاویر پل، حاوی برچسب‌های شکاف برای آموزش شبکه YOLOv4-FPM استفاده می‌کند. شبکه با استفاده از فرایند دو مرحله‌ای، ابتدا در یک مجموعه داده بزرگ آموزش می‌بیند و سپس در مجموعه داده شکاف پل، به طور دقیق تنظیم می‌شود. این فرایند به شبکه کمک می‌کند تا ویژگی‌های شکاف را یاد بگیرد و دقت تشخیص را بهبود بخشد. برای مدیریت تغییرات مقیاس شکاف‌های پل، ماژول Feature Pyramid به معماری YOLOv4 اضافه شده است. این ماژول ویژگی‌های چند مقیاس را از سطوح مختلف استخراج می‌کند و تشخیص شکاف‌ها در اندازه‌ها و رزولوشن‌های مختلف را ممکن می‌سازد. این ماژول توانایی شبکه در ثبت اطلاعات شکاف‌های محلی و سراسری را تقویت می‌کند [۱۰].

این پژوهش [۱۱] بر روی استفاده از شبکه‌های تولیدی مبارزه‌ای^۱ برای بهبود تشخیص علائم ترافیک تمرکز دارد. این مطالعه تصاویر مصنوعی را با تصاویر اصلی ترکیب کرده و از این روش برای افزایش مجموعه داده‌ها و ارزیابی اثر بخشی مجموعه داده‌های مصنوعی استفاده کرده است. پژوهشگران از تعداد و اندازه‌های مختلفی از تصاویر برای آموزش استفاده کرده‌اند و از شاخص شباهت ساختاری ($SSIM^2$) و خطای میانگین مربعات (MSE^3) برای ارزیابی کیفیت تصویر استفاده می‌کنند. پژوهشگران مجموعه داده تصاویر اصلی را با تصاویر مصنوعی گسترش داده و مدل تصویر اصلی را با مدل تصویر سنتز شده مقایسه کردند.

پس از ترکیب تصویر واقعی با تصویر سنتز شده توسط $LSGAN^4$ ، عملکرد تشخیص به صورت قابل ملاحظه‌ای بهبود یافت و دقت آن در YOLOv3 به ۸۴.۹٪ و در YOLOv4 به ۸۹.۳۳٪ رسید [۱۱].

این مقاله [۱۲] یک چارچوب تشخیص دهنده اشیا یک مرحله‌ای برای رانندگی خودکار ارائه می‌دهد که بر YOLOv4 استوار است. این چارچوب با استفاده از شبکه پایه $CSP^5Darknet53_dcn^6(P^7)$ ، دقت تشخیص را بهبود می‌بخشد. لایه آخر خروجی در $CSPDarknet53$ با کانولوشن تغییرپذیر جایگزین می‌شود تا دقت تشخیص را بهبود دهد. الگوریتم YOLOv4-5D باعث افزایش میانگین دقت تشخیص در مجموعه داده‌های BDD^۸ به میزان ۴.۲۳٪ و در مجموعه داده‌های KITTI^۹ به میزان ۱.۶۸٪ می‌شود. مدل سرعت استنتاج را ۳۱.۳٪ افزایش و استفاده از حافظه را به ۹۸.۱ مگابایت کاهش می‌دهد. الگوریتم پیشنهادی قادر به تشخیص با سرعت بیشتر از ۶۶ فریم در ثانیه است و دقت بالاتری نسبت به روش‌های قبلی ارائه می‌دهد [۱۲].

این مقاله [۱۳] یک شبکه عصبی کانولوشن عمیق را برای بهبود تشخیص علائم ترافیکی با استفاده از YOLOv4 ارائه می‌دهد. در این روش، شبکه عصبی عمیق برای تشخیص علامت‌های ترافیکی طراحی شده است که با استفاده از معماری مبتنی بر YOLOv4، این روش قادر است علامت‌های ترافیکی را با دقت و سرعت بالا تشخیص دهد.

¹ Generative Adversarial Networks

² Structural Similarity Index

³ Mean Squared Error

⁴ Least Squares Generative Adversarial Networks

⁵ Cross-Stage Partial connections

⁶ Deformable Convolutional Networks

⁷ Parallel processing

⁸ Behavior Driven Development

⁹ Karlsruhe Institute of Technology and Toyota Technological Institute

شبکه عصبی در این روش با استفاده از لایه‌های عمیق کانولوشنی و ادغام، ویژگی‌های مهم را استخراج می‌کند. سپس با استفاده از لایه‌های کاملاً متصل و لایه‌های تشخیص، علامت‌های ترافیکی را تشخیص می‌دهد. این روش از مزایای معماری YOLOv4 استفاده می‌کند که شامل عملکرد و سرعت بالا و قدرت تشخیص عالی است. با استفاده از این روش، دقت تشخیص علامت‌های ترافیکی به طور قابل توجهی افزایش می‌یابد. این شبکه عصبی عمیق، قابلیت یادگیری و بهبود پذیری دارد.

این مقاله [۱۴] یک مدل تشخیص شیء با جزئیات دقیق و سریع را بر اساس شبکه عصبی عمیق YOLOv4 ارائه می‌دهد. این مدل برای تشخیص شیء با جزئیات دقیق طراحی شده است و قادر است با دقت و سرعت بالا اقدام به تشخیص اشیا کند. در این مدل از شبکه عصبی عمیق YOLOv4 استفاده شده است تا با بهره‌گیری از لایه‌های کانولوشنی و ادغام، ویژگی‌های کلیدی و جزئیات دقیق اشیا استخراج شود. سپس لایه‌های خروجی مدل، عملیات تشخیص را با دقت و سرعت بالا انجام می‌دهند. به این ترتیب، مدل قادر است اشیا را با جزئیات کامل و در زمان کوتاه شناسایی کند [۱۴].

این مقاله یک چارچوب تشخیص بصری مبتنی بر یادگیری عمیق را با سرعت و کارایی بالا برای بازرسی اجزای راه‌آهن ارائه می‌دهد. شبکه عصبی کانولوشنی پیشرفته YOLOv4 بر اساس تصاویر موجود در پایگاه داده تصاویر، اجزای راه‌آهن را ارزیابی می‌کند. این شبکه با استفاده از لایه‌های کانولوشنی و لایه‌های دیگر، ویژگی‌های مهم اجزای ریل‌ها را استخراج کرده و با استفاده از لایه‌های مذکور، تشخیص در سریع‌ترین زمان را انجام می‌دهد. با استفاده از این روش، می‌توان به طور دقیق اجزای ریل‌های راه‌آهن مانند استوانه‌ها و اتصالات را تشخیص داد و به صورت هم‌زمان بررسی نمود. این روش از دقت و سرعت بالا برخوردار است و قابلیت تشخیص در شرایط واقعی را داراست. از این روش می‌توان در حوزه‌هایی مانند نگهداری و تعمیرات راه‌آهن، امنیت و بررسی اجزای ریل‌ها استفاده کرد [۱۵].

این مقاله [۱۶] یک مدل جدید برای تشخیص هدف در تصاویر پهپادی با نام YOLOv4_Drone ارائه می‌دهد که بر اساس نسخه بهبود یافته الگوریتم YOLO بنا شده است. نتایج آزمایشی نشان می‌دهند که مدل پیشنهادی برای تشخیص هدف در تصاویر پهپادی YOLOv4_Drone نسبت به الگوریتم YOLOv4 بهبود داشته است که نشانگر کارآمدی روش است.

۳- نسخه بهبود یافته YOLOv4 با استفاده از الگوریتم IASD^۱

مدل شبیه‌سازی شده این الگوریتم شامل فرایندهایی است که موجب می‌شود YOLOv4 از نظر دقت و سرعت در تشخیص به نسخه بهبود یافته تبدیل شود. YOLOv4 یک الگوریتم تشخیص اشیا در تصاویر و ویدئوها است که بر اساس شبکه عصبی عمیق ساخته شده است. در مرحله پیش‌پردازش تصاویر در YOLOv4، عملیاتی اعمال می‌شود تا تصاویر به صورتی مناسب برای ورود به شبکه عصبی آماده شوند. این عملیات شامل تغییر اندازه تصاویر، تبدیل آن‌ها به فضای رنگی مشخص، نرمال‌سازی و استانداردسازی داده‌ها و حذف نویز می‌شود.

پس از انجام عملیات پیش‌پردازش، تصاویر آماده به عنوان ورودی به شبکه عصبی YOLOv4 داده می‌شوند. در این مرحله، شبکه عصبی عمیق با استفاده از لایه‌های پیچشی^۲ و لایه‌های تماماً متصل^۳، تشخیص اشیا را در تصاویر انجام می‌دهد. پس از خروج نتایج از شبکه، کار تمام می‌شود؛ اما در این مقاله سعی بر این است که از پس‌پردازش برای بهبود و تفسیر نتایج استفاده شود.

^۱ Improving Accuracy and Speed in Detection

^۲ Convolutional Layers

^۳ Fully Connected Layers

خروجی YOLOv4 شامل مختصات و اطلاعات مربوط به اشیاء تشخیص داده شده است. این داده‌ها موقعیت مکانی مستطیلی شیء، میزان اطمینان از تشخیص و برجسب کلاس مربوطه را در بر می‌گیرند. با استفاده از پس‌پردازش، می‌توان این خروجی‌ها را تجزیه و تحلیل کرده و به شیوه‌های مختلفی نمایش داده و اطلاعات مورد نیاز را استخراج کرد [۱۷].

همچنین، پس‌پردازش در YOLOv4 می‌تواند شامل تنظیم پارامترهای شبکه باشد. این شامل تنظیماتی مانند آستانه‌ها برای تصمیم‌گیری در مورد وجود یا عدم وجود شیء، تنظیمات مربوط به تشخیص شیء در حالت‌های خاص و سایر تنظیمات مربوط به عملکرد و دقت شبکه است.

نقش پس‌پردازش در نسخه بهبود یافته YOLOv4، بهبود دقت در تشخیص است. YOLOv4 از یک شبکه عصبی کانولوشنی برای پیش‌بینی کادرهای محدود کننده اشیاء در یک تصویر استفاده می‌کند. با این حال، پیش‌بینی‌های اولیه شبکه عصبی ممکن است دقیق نباشند. پس‌پردازش با استفاده از تکنیک‌های مختلفی مانند فیلتر کردن و حدس زدن، دقت تشخیص را بهبود می‌بخشد.

الگوریتم IASD در پنج گام اجرا می‌شود:

گام ۱: خروجی YOLOv4 شامل چندین جعبه محدود کننده است که ممکن است تکراری، نویزی یا با اطمینان کم باشند.

گام ۲: ابتدا جعبه‌هایی که امتیاز اطمینان‌شان کمتر از یک آستانه مشخص (θ) است، حذف می‌شوند.

گام ۳: با استفاده از Non-Maximum Suppression، جعبه‌های همپوشان حذف شده و تنها جعبه‌ای که بالاترین امتیاز اطمینان را دارد باقی می‌ماند.

گام ۴: جعبه‌های باقی‌مانده که بسیار نزدیک هستند با یکدیگر ادغام می‌شوند، سپس ابعادشان باز تنظیم می‌گردد تا مرز شیء دقیق‌تر شود.

گام ۵: جعبه‌های نهایی معتبر، در لیست خروجی قرار می‌گیرند.

¹ Accuracy

شبه کد الگوریتم IASD

ورودی‌ها:

تکنیک پس پردازش.

predictions: یک ساختار حاوی Bounding Boxes پیش‌بینی شده، کلاس‌ها، و امتیازات.

ds: یک مجموعه داده حاوی Bounding Boxes و کلاس‌های واقعی.

خروجی:

Bounding Boxes: جعبه‌های فیلتر شده

classes: کلاس‌های فیلتر شده

scores: امتیازات فیلتر شده

۱ فیلتر کردن Bounding Boxes پیش‌بینی شده:

برای هر Bounding Box پیش‌بینی شده:

امتیاز آن را محاسبه کنید.

اگر

امتیاز Bounding Box برابر یا بیشتر از یک آستانه خاص باشد، آن را نگه دارید.

پایان

پایان

۲ اختصاص کلاس‌ها به Bounding Boxes فیلتر شده:

برای هر Bounding Box فیلتر شده:

کلاس با بیشترین احتمال را پیدا کنید.

احتمالات بر اساس امتیازات Bounding Boxes محاسبه می‌شوند.

پایان

حذف کادرهای محدود کننده ضعیف و تکراری: پس پردازش از اشتراک بر اجتماع^۱ برای حذف کادرهای محدود کننده ضعیف و تکراری استفاده می‌کند. کادرهای محدود کننده‌ای که اشتراک بر اجتماع آن‌ها کمتر از آستانه باشد، حذف می‌شوند. این امر به بهبود دقت تشخیص کمک می‌کند؛ زیرا کادرهای محدود کننده ضعیف و تکراری به احتمال زیاد برای شناسایی شیء اشتباه می‌کنند [۱۸].

بهبود دقت تشخیص اشیا کوچک و اشیا با همپوشانی: پس پردازش از الگوریتم‌های حدس زدن برای بهبود دقت تشخیص اشیا کوچک و اشیا با همپوشانی استفاده می‌کند. این الگوریتم‌ها از ویژگی‌های کادرهای محدود کننده و امتیازات شناسایی برای حدس زدن کلاس صحیح شیء استفاده می‌کنند. این امر به بهبود دقت تشخیص اشیا کوچک و اشیا با همپوشانی کمک می‌کند؛ زیرا این اشیا ممکن است به طور دقیق توسط شبکه عصبی اولیه شناسایی نشوند.

¹ Intersection over Union

کاهش نویز در خروجی الگوریتم: پس‌پردازش می‌تواند به کاهش نویز در خروجی الگوریتم کمک کند. این امر به بهبود دقت تشخیص کمک می‌کند؛ زیرا نویز می‌تواند منجر به شناسایی شیء اشتباه شود. در مجموع، پس‌پردازش نقش مهمی در بهبود دقت تشخیص YOLOv4 ایفا می‌کند.

۴- تنظیم الگوریتم‌های ADAM¹ و SGDM²

در مدل YOLOv4، الگوریتم‌های بهینه‌سازی مانند ADAM و SGDM می‌توانند برای آموزش شبکه عصبی استفاده شوند. این الگوریتم‌ها به منظور بهبود فرایند یادگیری و کاهش تابع هدف (تابع هزینه) مورد استفاده قرار می‌گیرند. در YOLOv4، این الگوریتم‌ها می‌توانند به عنوان بهینه‌سازی استفاده شوند تا شبکه عصبی در فرایند آموزش بهبود یابد. اما برای انتخاب مناسب‌ترین الگوریتم بهینه‌سازی و تنظیم پارامترهای آن، نیاز به ارزیابی و آزمایش بر روی مسئله و داده‌های خاصی داریم. پارامترهای الگوریتم بهینه‌سازی وابسته به مسئله، مجموعه داده و محدودیت‌های مربوطه باید انجام شود.

SGDM نیز یک الگوریتم بهینه‌سازی است که بر پایه گرادینان کاهشی تصادفی^۳ با عامل سرعت است. عامل سرعت به منظور حرکت سریع‌تر در فضای جستجو و از بین بردن اثر نویز و عدم قطعیت در گرادینان‌ها استفاده می‌شود. این عامل با افزودن یک مقدار تنظیمی به گرادینان‌ها، موجب کاهش نوسانات و جلوگیری از تغییرات ناخواسته در جهت گرادینان می‌شود. در SGDM نیز می‌توان نرخ یادگیری و عامل سرعت را تنظیم کرد. در این فرمول:

$$\begin{aligned} V(t) &= \alpha \times v(t-1) - \eta \times \nabla J(w(t-1)) \\ w(t) &= W(t-1) + V(t) \end{aligned} \quad (1)$$

نشان‌دهنده مقدار وزن‌های مورد استفاده در مرحله بعدی است. این مقدار با استفاده از مقدار وزن‌های قبلی و نرخ یادگیری (η) و گرادینان تابع هزینه ($\nabla J(w(t-1))$) محاسبه می‌شود.

α نشان‌دهنده ضریب تکانه است که معمولاً در بازه $[0, 1]$ قرار می‌گیرد. این ضریب مقداری بین ۰ و ۱ است که نشان می‌دهد چه میزان از وزن‌های قبلی باید در محاسبه وزن‌های جدید مؤثر باشند. η نرخ یادگیری است که معمولاً مقدار ثابتی در بازه $(0, 1)$ است. این مقدار نشان می‌دهد چقدر باید وزن‌ها در هر مرحله به سمت گرادینان کاهش یابند. $\nabla J(w(t-1))$ گرادینان تابع هزینه نسبت به وزن‌ها در مرحله قبلی است. $w(t)$ وزن‌های جدید در مرحله جاری هستند که با استفاده از مقدار وزن‌های قبلی و مقدار $v(t)$ محاسبه می‌شوند.

SGDM با استفاده از ضریب تکانه می‌تواند بهبود قابل توجهی در سرعت و کیفیت یادگیری شبکه‌های عصبی داشته باشد، زیرا از گرادینان‌های گذشته استفاده می‌کند تا به سرعت تغییرات و جهت بهینه را به شبکه برساند.

ADAM یک الگوریتم بهینه‌سازی است که برای تطبیق خودکار نرخ یادگیری با چندین پارامتر، کارآمد است. این الگوریتم با استفاده از گرادینان‌های انتقالی^۴ از طولانی‌مدت و میانگین مربعات گرادینان‌ها، نرخ یادگیری را تطبیق می‌دهد. ADAM دارای دو

¹ Adaptive Moment Estimation

² Stochastic Gradient Descent with Momentum

³ Stochastic Gradient Descent

⁴ Moving Average

پارامتر اصلی است که قابل تنظیم هستند: نرخ یادگیری و بتا^۱ که مقادیر بهینه آن‌ها وابسته به مسئله و مجموعه داده مورد استفاده قرار می‌گیرد.

$$m(t) = \beta_1 \times m(t-1) + (1 - \beta_1) \times \nabla J(w(t-1)) \quad (2)$$

$$v(t) = \beta_2 \times v(t-1) + (1 - \beta_2) \times (\nabla J(w(t-1)))^2 \quad (3)$$

$$\hat{m}(t) = \frac{m(t)}{(1 - \beta_1^t)} \quad (4)$$

$$\hat{v}(t) = \frac{v(t)}{(1 - \beta_2^t)} \quad (5)$$

$$w(t) = w(t-1) - \eta \times \frac{\hat{m}(t)}{(\sqrt{\hat{v}_1(t)} + E)} \quad (6)$$

$m(t)$ نشان‌دهنده یک تخمین اولیه از میانگین گرادیان است. این مقدار با استفاده از مقدار $m(t-1)$ و گرادیان تابع هزینه $(\nabla J(w(t-1)))$ (و ضرب β_1 محاسبه می‌شود. $v(t)$ نشان‌دهنده یک تخمین اولیه از میانگین مربعات گرادیان است. این مقدار با استفاده از مقدار $v(t-1)$ و گرادیان تابع هزینه $(\nabla J(w(t-1)))^2$ و ضرب β_2 محاسبه می‌شود. مقادیر $m(t)$ و $v(t)$ تخمینی از میانگین گرادیان و میانگین مربعات گرادیان‌اند که در ابتدای فرایند آموزش به منظور کاهش اثر تمایل $m(t)$ و $v(t)$ به سمت صفر به کار گرفته می‌شوند. β_1 و β_2 ضرایبی هستند که معمولاً در بازه $(0, 1)$ قرار می‌گیرند. این ضرایب نشان‌دهنده نسبت اهمیت گرادیان و مربعات گرادیان به تخمین‌های قبلی است. η نرخ یادگیری است که معمولاً مقدار ثابتی در بازه $(0, 1)$ است.

۴-۱- پارامترهای تنظیم‌کننده

نرخ یادگیری^۲ نشان‌دهنده میزان تغییراتی است که در هر مرحله به وزن‌ها اعمال می‌شود. مقدار این پارامتر باید به گونه‌ای تنظیم شود که بهبودی پیوسته در طول آموزش حاصل شود. در صورت انتخاب نرخ یادگیری بسیار بزرگ، الگوریتم ممکن است به جای همگرایی در نقطه بهینه، پیرامون آن نوسان کرده یا از مسیر بهینه منحرف شود.

اگر نرخ یادگیری بیش از حد کوچک باشد، ممکن است همگرایی به جواب بهینه بسیار کند شود یا حتی همگرایی به جواب بهینه اصلاً رخ ندهد؛ بنابراین، تنظیم مناسب نرخ یادگیری یکی از مهم‌ترین چالش‌ها در آموزش الگوریتم‌های یادگیری عمیق است.

اندازه دسته^۳ نشان‌دهنده تعداد نمونه‌ها است که در هر مرحله برای محاسبه گرادیان استفاده می‌شود. مقدار مناسب برای اندازه دسته باید به گونه‌ای انتخاب شود که مصرف منابع محاسباتی از جمله حافظه و توان پردازشی در سطحی بهینه و معقول حفظ شود.

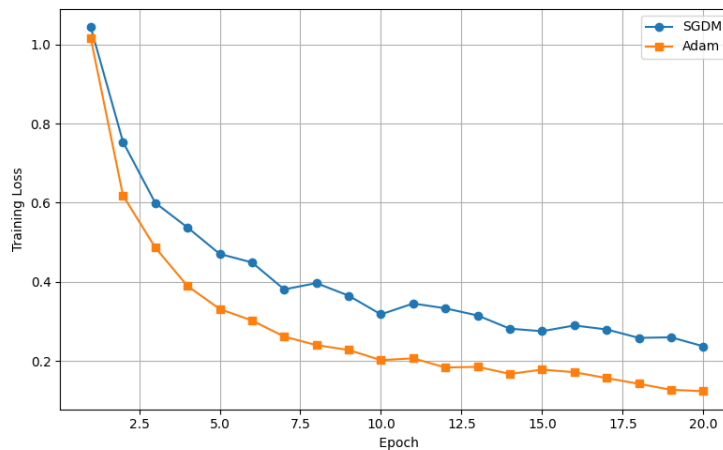
¹ Beta

² Learning Rate

³ Batch Size

اندازه دسته بزرگ‌تر معمولاً به معنای استفاده از بیشترین توان محاسباتی است و می‌تواند سرعت آموزش را افزایش دهد، اما ممکن است بهبود کیفی در نتایج دسته‌بندی را کاهش دهد. از طرف دیگر، اندازه دسته کوچک‌تر می‌تواند دقت را افزایش دهد، اما آموزش را به طور کلی کندتر کند.

حداکثر تعداد دوره‌های آموزشی^۱ تعیین می‌کند که فرایند آموزش تا چه مدت ادامه یابد. این پارامتر معمولاً بر اساس تعداد دوره‌هایی تنظیم می‌شود که طی آن، بهبود قابل توجهی در عملکرد مدل مشاهده می‌گردد. در مراحل ابتدایی، معمولاً از تعداد دوره‌های کمتر استفاده می‌شود و سپس با ارزیابی روند پیشرفت مدل، این مقدار به گونه‌ای تنظیم می‌گردد که بهبود تدریجی و همگرایی به سمت جواب بهینه حاصل شود. با این حال، افزایش بیش از حد تعداد دوره‌های آموزشی می‌تواند منجر به بروز پدیده‌ی بیش‌برازش شود. در این حالت، مدل به جای یادگیری الگوهای کلی داده‌ها، ویژگی‌های مجموعه‌ی آموزشی را به صورت خاص حفظ می‌کند و در نتیجه تنها بر روی همان داده‌ها عملکرد خوبی دارد، اما در مجموعه‌های جدید دقت خود را از دست می‌دهد.



شکل (۱) مقایسه روند کاهش خطا بین دو الگوریتم بهینه‌ساز SGDM و ADAM

شکل (۱) روند کاهش خطا^۲ را در طول دوره‌های آموزشی برای دو الگوریتم بهینه‌ساز SGDM و ADAM نشان می‌دهد. همان‌طور که دیده می‌شود، SGDM گرچه روندی پایدار دارد، اما سرعت همگرایی آن کمتر از ADAM است. در مقابل، ADAM در مراحل اولیه با سرعت بیشتری بهینه می‌شود، هر چند ممکن است نوسانات بیشتری داشته باشد. ترکیب این دو الگوریتم در پژوهش حاضر باعث شده است که هم سرعت همگرایی و هم پایداری آموزش بهبود یابد.

¹ Maximum Number of Training Epochs

² Training Loss

شبه کد الگوریتم بهینه‌سازی ترکیبی ADAM و SGDM

ورودی‌ها:

مجموعه داده‌های آموزشی: مجموعه داده، شامل نمونه‌هایی است که مدل بر اساس آن‌ها آموزش می‌بیند.

۱ تنظیم نرخ یادگیری:

برای افزایش دقت:

نرخ یادگیری را کاهش می‌دهیم.

اگر

افزایش سرعت تشخیص را می‌خواهیم:

نیازمند بالا بردن نرخ یادگیری است

پایان

پایان

۲ تنظیم اندازه دسته:

برای افزایش دقت:

اندازه دسته را کاهش می‌دهیم.

اگر

افزایش سرعت تشخیص را می‌خواهیم:

نیازمند بالا بردن اندازه دسته است

پایان

پایان

۳ تنظیم حداکثر تعداد دوره‌های آموزشی:

برای افزایش دقت:

حداکثر تعداد دوره‌های آموزشی را افزایش می‌دهیم.

اگر

افزایش سرعت تشخیص را می‌خواهیم:

نیازمند کاهش حداکثر تعداد دوره‌های آموزشی است

پایان

پایان

۴-۲- تشخیص دهنده و شبکه عصبی مورد استفاده

با توجه به داده‌های جدول (۱) تشخیص دهنده با استفاده از شبکه عصبی و جعبه‌های کدگذار، قادر به تشخیص اشیا موجود در تصاویر است. جعبه‌های کدگذار یک مجموعه از مستطیل‌های ثابت هستند که برای تشخیص اشیا به کار می‌روند. هر جعبه کدگذار شامل اطلاعات مکانی و ابعادی است که به شبکه عصبی کمک می‌کند تا اشیا را در تصویر تشخیص دهد. نام‌های کلاس نیز مربوط به

دسته‌بندی مختلف اشیاء هستند که تشخیص دهنده می‌تواند تشخیص دهد. این نام‌ها معمولاً شامل اشیاء معروف مانند ماشین و غیره هستند.

جدول (۱) مشخصات تشخیص دهنده

شبکه عصبی	[1×1 dl network]
اجعه‌های پایه	[2×1 cell]
نام‌های کلاس	[80×1 cell]
اندازه ورودی	[416 416 3]
نام مدل	tiny-yolov4-coco

با ورودی‌هایی به اندازه [۴۱۶ ۴۱۶ ۳]، این آشکارساز قادر است اشیاء را در تصاویر تشخیص داده و نتایج را ارائه کند. با استفاده از این تشخیص دهنده و ویژگی‌ها، می‌توانید تصاویر را به عنوان ورودی داده و نتایج تشخیص که شامل مکان اشیاء و برجسب آن‌ها را دریافت کنید. در این مثال، آشکارساز، تشخیص وسایل نقلیه از تصاویر را انجام می‌دهد.

نتایج دقیق و کیفیت آن‌ها بستگی به عواملی مانند دقت مدل، تنظیمات پیش‌پردازش، کیفیت تصویر و مقدار داده آموزشی دارد. همچنین، مدل‌های مختلف می‌توانند دقت و نتایج متفاوتی داشته باشند. معمولاً مدل‌هایی با دقت بالا و آموزش بر روی مجموعه داده‌های گسترده، نتایج بهتری ارائه می‌دهند.

شبکه عصبی شامل لایه CNN با ابعاد [۱×۷۴ n net CNN Layer] است. با توجه به جدول (۲) این لایه شامل ۷۴ فیلتر با اندازه $n \times 1$ است که برای استخراج ویژگی‌های مختلف از ورودی استفاده می‌شوند.

جدول (۲) مشخصات شبکه عصبی عمیق

لایه‌ها	[74×1 n net CNN Layer]
اتصالات	[80×2 table]
پارامترهای قابل یادگیری ^۲	[80×3 table]
وضعیت	[38×3 table]
نام‌های ورودی	['input_1']
نام‌های خروجی	['conv_31' 'conv_38']
مقداردهی اولیه	1

شبکه عصبی دارای اتصالاتی با ابعاد [table ۲×۸۰] است. این اتصالات نشان‌دهنده ارتباطات بین لایه‌های مختلف شبکه و نحوه جریان اطلاعات است.

¹ Anchor Boxes

² Learnable

شبکه عصبی دارای پارامترهای قابل یادگیری با ابعاد [table ۳×۸۰] است. این پارامترها معمولاً به صورت وزن‌ها و بایاس‌ها^۱ بین لایه‌ها استفاده می‌شوند و در هنگام آموزش شبکه تغییر می‌کنند. در شبکه عصبی، جدولی با ابعاد [۳×۳۸] برای نمایش وضعیت وزن‌ها و پارامترهای داخلی در هر لحظه در نظر گرفته شده است. ورودی شبکه عصبی با نام 'input_1' و خروجی با نام‌های 'conv_31' و 'conv_38' تعیین شده است. این نام‌ها معمولاً برای ارجاع به لایه‌ها و تصاویر میانی در شبکه استفاده می‌شوند.

۵- نتایج شبیه‌سازی

در این مقاله، به بهبود دقت و سرعت YOLOv4 تمرکز شده است. با استفاده از تکنیک پس‌پردازش بر روی خروجی شبکه و انجام عملیاتی مانند حذف داده‌های تکراری و تصحیح موقعیت اشیاء، دقت تشخیص YOLOv4 بهبود یافت. علاوه بر این، با بهینه‌سازی نرخ یادگیری و تنظیم مناسب آن با استفاده از الگوریتم‌های بهینه‌سازی مانند ADAM یا SGDM، می‌توان عملکرد آموزشی YOLOv4 را بهبود بخشید. همچنین، با تنظیم مناسب اندازه دسته و تعیین حداکثر تعداد دوره‌های آموزشی، می‌توان سرعت آموزش YOLOv4 را افزایش داد و به عملکرد بهتری دست یافت.

همان‌گونه که جدول (۳) نشان می‌دهد، نسخه‌های جدید YOLO، در برخی معیارها از YOLOv4 عملکرد بهتری دارند، به‌ویژه در mAP^1 (IoU=0.5:0.95) و سرعت پردازش بر پایه GPU³. با این حال، الگوریتم پیشنهادی (YOLOv4+IASD) توانسته است نسبت به YOLOv4 اصلی، هم در mAP (IoU=0.5) و هم در mAP (IoU=0.5:0.95) بهبود قابل توجهی داشته باشد، بدون آنکه نیازمند منابع سخت‌افزاری سنگین باشد. این مزیت برای کاربردهای بلادرنگ در سامانه‌های کم‌منبع مانند پهپادها یا دوربین‌های نظارتی تعبیه‌شده بسیار کلیدی است، چرا که نسخه‌های جدید YOLO عمدتاً برای پردازنده‌های گرافیکی قدرتمند طراحی شده‌اند.

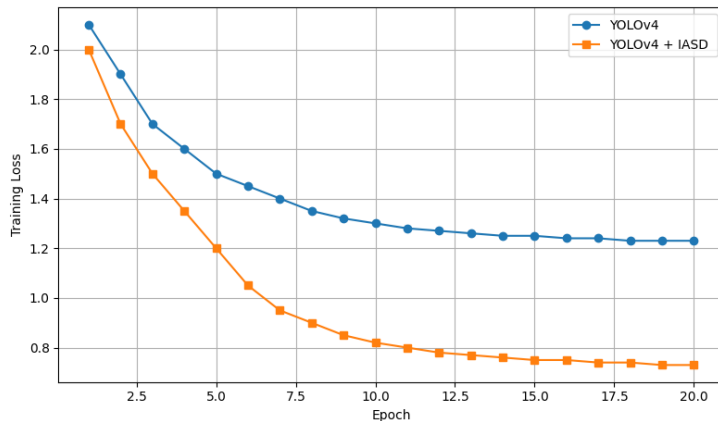
جدول (۳) مقایسه روش پیشنهادی با نسخه‌های جدید YOLO

نیاز به GPU/CPU	FPS	mAP (IoU=0.5:0.95) (%)	mAP (IoU=0.5) (%)	الگوریتم
متوسط	65	66.0	89.3	YOLOv4
بالا (GPU)	140	69.5	91.2	YOLOv5
بالا (GPU)	160	71.3	92.5	YOLOv7
بالا (GPU)	180	72.4	93.1	YOLOv8
کم (CPU هم قابل اجرا)	70	68.4	90.1	YOLOv4 + IASD

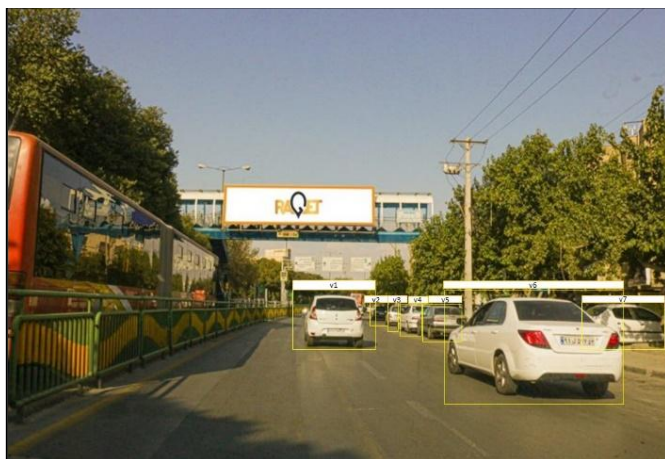
شکل (۲) روند کاهش تابع هزینه را در طول ۲۰ دوره آموزشی برای دو مدل YOLOv4 اصلی و نسخه پیشنهادی (YOLOv4+IASD) نشان می‌دهد. همان‌طور که مشاهده می‌شود، مدل پیشنهادی در مراحل اولیه آموزش با سرعت بیشتری به

¹ Bias

همگرایی می‌رسد و مقدار نهایی تابع هزینه نیز در آن پایین‌تر از YOLOv4 اصلی است. این موضوع بیانگر آن است که ترکیب بهینه‌سازهای ADAM و SGDM به همراه تابع پدازش IASD موجب بهبود کارایی فرایند آموزش و در نهایت افزایش دقت مدل در مرحله آزمون شده است.



شکل (۲) روند کاهش تابع هزینه در طول ۲۰ دوره آموزشی برای دو مدل YOLOv4 اصلی و نسخه پیشنهادی



شکل (۳) تشخیص وسایل نقلیه با استفاده از الگوریتم بهبود یافته YOLOv4



شکل (۴) تشخیص وسایل نقلیه با استفاده از الگوریتم اصلی YOLOv4

جدول (۴) پارامترهای کلیدی مورد استفاده در آموزش مدل پیشنهادی را نشان می‌دهد. انتخاب این مقادیر بر اساس مطالعات پیشین و آزمایش‌های تجربی صورت گرفته است تا تعادلی میان سرعت همگرایی، پایداری آموزش و دقت نهایی حاصل شود. به‌ویژه، استفاده هم‌زمان از دو بهینه‌ساز ADAM و SGDM با نرخ یادگیری مناسب و ضریب تکانه بهینه، باعث کاهش نوسانات گرادیان و دستیابی به دقت بالاتر را در مدت آموزش کوتاه‌تر فراهم کرده است.

جدول (۴) پارامترهای کلیدی مورد استفاده در آموزش مدل پیشنهادی

توضیح	مقدار انتخاب شده	پارامتر
مقدار اولیه برای شروع آموزش	0.001	نرخ یادگیری
تعادل بین سرعت و مصرف حافظه	32	اندازه دسته
تأسیس به همگرایی مدل	50	تعداد دوره
ضریب تکانه ^۴ در ADAM	0.9	β_1 (ADAM)
ضریب کاهش واریانس گرادیان	0.999	β_2 (ADAM)
سرعت همگرایی در SGDM	0.9	تکانه (SGDM)
برای افزایش دقت و سرعت هم‌زمان	ADAM + SGDM	بهینه‌سازی ترکیبی
مطابق استاندارد YOLOv4	416×416	اندازه ورودی تصاویر
سنجش جامع دقت و سرعت	FPS, mAP (IoU=0.5:0.95), mAP (IoU=0.5)	معیارهای ارزیابی

در این پژوهش یک نسخه بهبود یافته از الگوریتم YOLOv4 برای تشخیص وسایل نقلیه معرفی شد. نوآوری اصلی در ترکیب دو بهینه‌ساز ADAM و SGDM و نیز ارائه یک تابع پدازش سبک‌وزن با عنوان IASD بود که موجب بهبود هم‌زمان دقت و سرعت تشخیص گردید. نتایج تجربی نشان داد که روش پیشنهادی در مقایسه با YOLOv4 اصلی، هم در معیار mAP (IoU=0.5) و هم در معیار mAP (IoU=0.5:0.95) بهبود قابل توجهی دارد و از نظر سرعت آموزش نیز سریع‌تر به همگرایی می‌رسد. همچنین مقایسه با نسخه‌های جدیدتر YOLO نشان داد که هرچند آن‌ها از نظر سخت‌افزاری قدرتمندتر عمل می‌کنند، اما الگوریتم پیشنهادی به دلیل نیاز کمتر به منابع محاسباتی، گزینه‌ای مناسب برای سامانه‌های تعبیه‌شده و کاربردهای بلادرنگ مانند پهپادها، سیستم‌های نظارت شهری و خودروهای هوشمند است. در نهایت می‌توان گفت روش ارائه‌شده پلی میان سادگی و کارایی فراهم کرده و مسیر را برای تحقیقات آینده در جهت بهبود معماری‌های سبک‌وزن و بهینه‌سازی الگوریتم‌های تشخیص شیء هموار می‌سازد.

۶- نتیجه‌گیری

در این پژوهش یک نسخه بهبود یافته از الگوریتم YOLOv4 برای تشخیص وسایل نقلیه ارائه گردید. نوآوری اصلی کار در دو محور متمرکز بود: نخست، طراحی یک الگوریتم پدازش سبک‌وزن با عنوان IASD که قادر است جعبه‌های تکراری و نویزی

را حذف کرده و دقت نهایی مدل را به‌طور محسوسی افزایش دهد؛ دوم، بهینه‌سازهای ADAM و SGDM به‌صورت ترکیبی استفاده شده‌اند تا با بهره‌گیری از مزایای هر دو، همگرایی مدل سریع‌تر و پایدارتر گردد. نتایج آزمایش‌ها بر روی مجموعه داده گسترده نشان داد که روش پیشنهادی، در مقایسه با YOLOv4 اصلی، هم از نظر میانگین دقت (mAP) و هم از نظر سرعت تشخیص (FPS) بهبود چشمگیری داشته است. همچنین نمودار روند کاهش خطا نشان داد که مدل پیشنهادی در مراحل اولیه آموزش با سرعت بیشتری به سمت همگرایی حرکت می‌کند. علاوه بر این، مقایسه با نسخه‌های جدیدتر YOLO بیانگر آن است که هرچند آن‌ها در برخی معیارها عملکرد بهتری دارند، اما روش پیشنهادی به دلیل نیاز سخت‌افزاری کمتر و قابلیت پیاده‌سازی بر روی سامانه‌های کم‌منبع، مزیتی عملی و قابل توجه برای کاربردهای بلادرنگ نظیر پهپادها، سیستم‌های نظارت شهری و سامانه‌های هوشمند حمل‌ونقل دارد. بدین ترتیب، این پژوهش پلی میان دقت بالا و سادگی محاسباتی فراهم می‌آورد. به‌عنوان چشم‌انداز آینده، توسعه و ارزیابی الگوریتم IASD در کنار سایر نسخه‌های جدید YOLO و همچنین آزمایش بر روی مجموعه‌داده‌های متنوع‌تر می‌تواند به ارتقای بیشتر دقت و پایداری در سامانه‌های هوشمند منجر شود.

مراجع

- [1] E. Ong et al. "Video Object Segmentation." Synthesis Lectures on Computer Vision (2024). <https://doi.org/10.4018/978-1-59904-845-1.CH106>.
- [2] Myeongjae Jang et al. "Zero and Narrow-Width Value-Aware Compression for Quantized Convolutional Neural Networks." IEEE Transactions on Computers, 73 (2024): 249-262. <https://doi.org/10.1109/TC.2023.3315051>.
- [3] Miaomiao Liu et al. "An Improved Adam Optimization Algorithm Combining Adaptive Coefficients and Composite Gradients Based on Randomized Block Coordinate Descent." Computational Intelligence and Neuroscience, (2023). <https://doi.org/10.1155/2023/4765891>.
- [4] Gonzales, S. "Enhancing Image Classification Performance: A Comparative Analysis of Optimization Algorithms." International Journal of Advanced Research in Science, Communication and Technology, 2023.
- [5] Zheng, Z., Zhang, T., Liu, Z., Liu, B., Li, Y., & Sun, C. "Arbitrarily Oriented Object Detection in Remote Sensing Images Based on Improved YOLOv4-CSP." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 9355-9368, 2022.
- [6] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. "YOLOv4: Optimal Speed and Accuracy of Object Detection." Submitted, 23 April 2020.
- [7] Yu, J., & Zhang, W. "Face Mask Wearing Detection Algorithm Based on Improved YOLO v4." Sensors, vol. 21, no. 9, pp. 3263, 2021. <https://doi.org/10.3390/s21093263>
- [8] Jiang, Z., Zhao, L., Li, S., & Jia, Y. "Real-time Object Detection Method Based on Improved YOLOv4-Tiny." Submitted, 9 November 2020 (v1), last revised 2 December 2020 (v2).
- [9] H. Sayyadi, M. Mohassel Fegghi, M. Atashbar, M. Nangir, and J. Sayyadi, "Efficient Four-Dimensional Routing Planning of Heterogeneous Drones Utilizing the Fuzzy Clustering Algorithm in Complex Entropy Environments," Tabriz Journal of Electrical Engineering, 2024. [Online]. Available: <https://doi.org/10.22034/tjee.2024.60800.4816>.

- [10] Zhenwei Yu, Yonggang Shen, Chenkai Shen, A real-time detection approach for bridge cracks based on YOLOv4-FPM, *Automation in Construction*, Volume 122, 2021.
- [11] C. Dewi, R. -C. Chen, Y. -T. Liu, X. Jiang and K. D. Hartomo, "Yolo V4 for Advanced Traffic Sign Recognition With Synthetic Training Data Generated by Various GAN," in *IEEE Access*, vol. 9, pp. 97228-97242, 2021, doi: 10.1109/ACCESS.2021.3094201.
- [12] Y. Cai et al., "YOLOv4-5D: An Effective and Efficient Object Detector for Autonomous Driving," in *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-13, 2021, Art no. 4503613, doi: 10.1109/TIM.2021.3065438.
- [13] Dewi, C., Chen, RC., Jiang, X. et al. Deep convolutional neural network for enhancing traffic sign recognition developed on Yolo V4. *Multimed Tools Appl* 81, 37821–37845 (2022). <https://doi.org/10.1007/s11042-022-12962-5>.
- [14] Roy, A.M., Bose, R. & Bhaduri, J. A fast accurate fine-grain object detection model based on YOLOv4 deep neural network. *Neural Comput & Applic* 34, 3895–3921 (2022). <https://doi.org/10.1007/s00521-021-06651-x>.
- [15] Feng Guo, Yu Qian, Yuefeng Shi, Real-time railroad track components inspection based on the improved YOLOv4 framework, 2021, (<https://www.sciencedirect.com/science/article/pii/S0926580521000479>).
- [16] Li Tan, Xinyue Lv, Xiaofeng Lian, Ge Wang, YOLOv4_Drone: UAV image target detection based on an improved YOLOv4 algorithm, *Computers & Electrical Engineering*, Volume 93, 2021.
- [17] Gibson, I., Rosen, D., Stucker, B., Khorasani, M. (2021). Post-Processing. In: *Additive Manufacturing Technologies*. Springer, Cham. https://doi.org/10.1007/978-3-030-56127-7_16.
- [18] H. Sayyadi, M. M. Feghhi, M. Nangir and J. Sayyadi, "Improving UAV-Based Monitoring of Solar Power Plants Using Coverage Path Planning Model with Adaptive Learning Algorithm," *2025 10th International Conference on Technology and Energy Management (ICTEM)*, Tabriz, Iran, Islamic Republic of, 2025, pp. 1-7, doi: 10.1109/ICTEM66196.2025.11063624.
- [19] J. Sayyadi, M. Nangir, M. M. Feghhi and H. Sayyadi, "Optimal Management of Solar Energy Generation Using Derivatives of Irradiation Angle Data with the Gradient Boosting Algorithm," *2025 10th International Conference on Technology and Energy Management (ICTEM)*, Tabriz, Iran, Islamic Republic of, 2025, pp. 1-6, doi: 10.1109/ICTEM66196.2025.11063646.



Sahand University
of Technology

DOR:

[20.1001.1.23223146.1404.12.1.6.9](https://doi.org/10.1001.1.23223146.1404.12.1.6.9)

Journal of Nonlinear
Systems in Electrical
Engineering

Vol.12, No.1

Spring and Summer 2025

ISSN: 2322 – 3146

<http://jnsee.sut.ac.ir>

Adaptive Model Predictive Control for Load Frequency Control Problem of Multi-Area Interconnected Power Systems with Time-Varying Parameters

Amin Khadem, Mohsen Ahmadian*, Ahmad Hajipour

Faculty of Electrical and Computer Engineering, Hakim Sabzevari University,
Sabzevar, Iran

energypak4271@chmail.ir; ahmadniamohsen85@gmail.com;
ahmad.hajipour@gmail.com

Abstract

With the rapid development of power generation technology and the increasing demand from power users, interconnected multi-area power systems have become a development trend. In this paper, an adaptive model predictive control scheme is proposed for the load frequency problem of an interconnected multi-area power system in the presence of parameter variations. The controller includes an autoregressive with exogenous inputs (ARX) model, a linear time-varying Kalman filter (LTVKF), and a recurrent polynomial model estimator (RPME) algorithm. A reduced-order ARX model of the original power system is identified using the recurrent polynomial model estimator and is used as the system prediction model in the adaptive model predictive control. The system state variables are observed using a linear time-varying Kalman filter, and the constrained receding horizon optimization problem is solved by quadratic programming. In addition to estimating a reduced-order model, the future control sequence is estimated using Laguerre basis functions to reduce the computational burden. Numerical simulations show that the proposed adaptive model predictive control is more efficient in rejecting load disturbances in power systems with time-varying parameters compared to other control schemes.

Keywords

Load frequency control; Adaptive model predictive control; Recursive polynomial model estimator; Autoregressive with exogenous inputs model; Linear time-varying Kalman filter.

1. Short Introduction

Load-frequency control (LFC) is critical in multi-area power systems, but high renewable penetration reduces inertia and challenges conventional PI/PID controllers. Model predictive control (MPC) improves performance and handles constraints, yet standard MPC faces high computational load and limited robustness. This paper proposes an adaptive MPC (AMPC) using a reduced-order data-driven model and time-varying Kalman filter to enhance robustness and reduce computation. The approach is validated on a three-area thermal system under load disturbances and generation constraints, outperforming standard MPC and PID control.

2. Proposed Work

In this paper, an adaptive model predictive control (AMPC) strategy is proposed for load-frequency control in a multi-area interconnected power system, effectively addressing both parametric variations and load disturbances. Initially, a reduced-order ARX model of the primary power system is identified offline using a recursive polynomial model estimator, which captures the essential dynamic characteristics of each control area while substantially reducing computational complexity. This simplified model enables efficient prediction of future system behavior, maintaining sufficient fidelity for accurate control design. In each sampling instant, the predictive controller updates the model online, allowing continuous adaptation to real-time variations in system parameters, generation conditions, and external disturbances. A time-varying Kalman filter (LTVKF) is integrated to provide precise estimation of both measurable and unmeasurable system states, mitigating the effects of sensor noise and unmodeled dynamics. The AMPC controller explicitly enforces input constraints and rate-of-change limits while regulating frequency deviations and tie-line power flows across multiple areas, ensuring safe and stable operation under varying load conditions. To achieve this, the control law leverages Laguerre functions to efficiently compute future control sequences, significantly reducing online computational burden compared to standard MPC. The proposed approach is particularly suitable for modern power grids with high penetration of renewable energy sources, where system uncertainties and load fluctuations are frequent. Numerical simulations are performed on a three-area interconnected power system under two scenarios: nominal parameters and parametric variations of up to 20% in generator and turbine time constants, inertia, and damping coefficients. The AMPC performance is compared against standard MPC and conventional PID controllers. Results demonstrate that the proposed AMPC substantially reduces overshoot, settling time, and control errors for all monitored variables, while providing smoother and faster dynamic responses even under continuous disturbances and parametric uncertainties. Moreover, the controller exhibits strong robustness to variations in system parameters, maintaining superior frequency regulation and tie-line power control. Computationally, AMPC achieves near-PID efficiency while delivering advanced predictive capabilities. These findings highlight that the proposed adaptive MPC not only improves dynamic performance and stability in multi-area power systems but also offers a practical and reliable alternative to traditional PID controllers, especially for modern power networks facing uncertain and time-varying operating conditions. The proposed methodology can also be extended to larger-scale multi-area networks, providing a scalable and adaptable framework for future smart grid applications, including automated load management and integration of distributed energy resources. Additionally, the adaptability of the controller ensures long-term reliability and resilience, crucial for maintaining power quality in the presence of unexpected disturbances or component failures. Furthermore, the AMPC framework facilitates coordination between control areas, enhancing system-wide stability and minimizing inter-area oscillations. The approach also supports future integration with real-time monitoring and predictive analytics, enabling proactive control actions in response to forecasted load or generation variations. Overall, this study establishes a comprehensive strategy for adaptive, high-performance, and resilient load-frequency control in modern interconnected power systems.

3. Conclusion

An adaptive MPC was proposed for LFC in multi-area interconnected power systems under parameter variations. A reduced-order ARX model was identified offline using a recursive polynomial estimator and updated online at each step, while a linear time-varying Kalman filter estimates system states. Compared to standard MPC and PID control, the proposed AMPC reduces computational burden and enhances dynamic performance, including faster settling time, lower overshoot, and improved disturbance rejection. Results demonstrate that AMPC is a robust, efficient, and practical alternative to conventional PID control for modern complex power systems.



دانشگاه صنعتی سهند

DOR:

20.1001.1.23223146.1404.12.1.6.9

نشریه علمی غیرنخبی در مهندسی برق

دوره ۱۲ - شماره ۱

بهار و تابستان ۱۴۰۴

صفحات ۱۱۸ الی ۱۳۹

ISSN: 2322-3146

http://jnsee.sut.ac.ir

کنترل پیش‌بین مدل تطبیقی برای مساله کنترل بار-فرکانس سیستم‌های قدرت بهم‌پیوسته چند ناحیه‌ای با پارامترهای متغیر با زمان

امین خادم^۱، محسن احمدنیا^{۲*} و احمد حاجی‌پور^۳

^۱کارشناس ارشد، دانشکده مهندسی برق و کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، energypak4271@chmail.ir

^۲نویسنده مسئول، استادیار، دانشکده مهندسی برق و کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، ahmadniamohsen85@gmail.com

^۳دانشیار، دانشکده مهندسی برق و کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، ahmad.hajipoor@gmail.com

تاریخ پذیرش: ۱۴۰۴/۸/۲۵

تاریخ بازنگری: ۱۴۰۴/۶/۳۰ و ۱۴۰۴/۸/۱۲

تاریخ ارسال: ۱۴۰۴/۴/۲۷

چکیده

با پیشرفت سریع فناوری تولید برق و افزایش تقاضا از سوی کاربران، سیستم‌های قدرت چند ناحیه‌ای بهم‌پیوسته به یک روند توسعه تبدیل شده‌اند. در این مقاله، یک طرح کنترل پیش‌بین مدل تطبیقی برای کنترل بار-فرکانس (LFC) یک سیستم قدرت چند ناحیه‌ای بهم‌پیوسته در حضور تغییرات پارامتری پیشنهاد شده است. این کنترل کننده شامل یک مدل خود رگرسیون با ورودی خارجی (ARX)، یک فیلتر کالمن خطی متغیر با زمان (LTVKF) و یک الگوریتم تخمینگر مدل چندجمله‌ای بازگشتی (RPME) می‌شود. با استفاده از تخمینگر مدل چندجمله‌ای بازگشتی یک مدل ARX با مرتبه کاهش یافته از سیستم قدرت اصلی شناسایی می‌شود و به عنوان مدل پیش‌بینی سیستم در کنترل پیش‌بین مدل تطبیقی مورد استفاده قرار می‌گیرد. متغیرهای حالت سیستم با کمک فیلتر کالمن خطی متغیر با زمان رویت می‌شوند و مساله بهینه‌سازی با افق متحرک مقید از طریق برنامه‌ریزی مربعی حل می‌شود. علاوه بر تخمین حاصل از مدل دارای کاهش مرتبه، دنباله کنترلی آینده با استفاده از توابع پایه لاگر تخمین زده می‌شود تا بار محاسباتی کاهش یابد. شبیه‌سازی‌های عددی نشان می‌دهند که کنترل پیش‌بین مدل تطبیقی پیشنهادی در مقایسه با سایر طرح‌های کنترلی، در دفع اغتشاشات بار در سیستم‌های قدرت با پارامترهای متغیر با زمان، کارآمدتر است.

واژه‌های کلیدی

کنترل بار-فرکانس،
کنترل پیش‌بین مدل تطبیقی،
تخمینگر مدل چندجمله‌ای بازگشتی،
مدل خود رگرسیون با ورودی
خارجی،
فیلتر کالمن خطی متغیر با زمان.

۱- مقدمه

به منظور حفظ فرکانس سیستم قدرت چند ناحیه‌ای بهم‌پیوسته در مقدار نامی و توان خطوط ارتباطی بین نواحی مجاور در مقادیر برنامه‌ریزی شده از کنترل بار-فرکانس^۱ (LFC) بهره گرفته می‌شود [۱]. به‌طور مرسوم، کنترل تولید خودکار برای حصول این اهداف بکار می‌رود به طوری که یک سیستم کنترلی در هر ناحیه، وظیفه ارسال سیگنال‌های لازم برای تنظیم بار مرجع ژنراتورهای درون ناحیه را بر عهده دارد. اما نفوذ رو به رشد واحدهای تولید تجدیدپذیر از قبیل مزارع بادی و خورشیدی، به ویژه واحدهای بر مبنای واسط مبدل فرکانسی، سبب ایجاد چالش در کنترل فرکانس سیستم قدرت می‌شود. از یک سو، مشخصه نوسانی و قابلیت پیش‌بینی محدود واحدهای تجدیدپذیر، کارایی سیستم کنترل تولید خودکار را با چالش همراه می‌سازد. از سوی دیگر، واحدهای تجدیدپذیر با واسط فرکانسی نقشی در تامین لختی چرخشی سیستم قدرت ندارند [۲]. هنگامی که ژنراتورهای سنکرون بیشتری با واحدهای تجدیدپذیر با واسط فرکانسی جایگزین شوند، لختی چرخشی سیستم به‌طور چشمگیری کاهش می‌یابد و حتی با زمان تغییر می‌کند که باعث می‌شود سیستم در برابر اغتشاشات، شکننده‌تر و کنترل آن دشوارتر شود [۳].

به دلیل طراحی ساده، قابلیت اطمینان و نسبت قیمت به عملکرد جذاب، مهندسان همچنان کنترل کننده تناسبی-انتگرالی [۴] و تناسبی-انتگرالی-مشتقی^۲ (PID) [۵] را برای کنترل بار-فرکانس در سیستم‌های قدرت واقعی انتخاب می‌کنند. با توجه به پیچیدگی رو به رشد سیستم‌های قدرت فعلی و چالش‌های موجود، کنترل‌های کلاسیک مذکور با ساختار ساده نمی‌توانند جوابگوی نیازهای کنترلی باشند و نیاز به روش‌های مدرن بیش از پیش احساس می‌شود. طراحی کنترل کننده‌های بر مبنای PID با درجات آزادی بیشتر می‌تواند تا حد زیادی محدودیت‌های کنترل‌های مرسوم را برطرف کند [۶]. علاوه بر آن، در سال‌های اخیر، روش‌های کنترلی مدرن برای طراحی کنترل کننده بار-فرکانس معرفی شده است که از این میان می‌توان به کنترل مد لغزشی [۷]، کنترل پیش‌بین مدل^۳ (MPC) [۳]، کنترل منطق فازی [۸]، کنترل شبکه عصبی مصنوعی [۹]، یادگیری تقویتی [۱۰]، کنترل H_∞ [۸ و ۱۱]، کنترل بر مبنای ناتساوی ماتریسی خطی [۱۲]، کنترل بر مبنای رویکر اغتشاش غیرخطی [۱۳]، رگولاتور مربعی خطی [۱۴] و کنترل تطبیقی بر مبنای فیلتر کالمن [۱۴-۱۵] اشاره کرد.

کنترل پیش‌بین مدل که یک طرح کنترل مدرن بر مبنای مدل شناخته می‌شود، دارای مشخصه‌هایی از قبیل پاسخ بهینه، قابلیت اعمال به سیستم‌های چند-ورودی چند-خروجی و جایدهی خودکار قیود در روند محاسبه عمل کنترلی است. در این روش، یک مساله بهینه‌سازی آنلاین با افق محدود با استفاده از روش برنامه‌ریزی مربعی برای محاسبه عمل کنترلی در هر پله زمانی حل می‌شود. کنترل پیش‌بین مدل برای مساله کنترل بار-فرکانس برای اولین بار در [۱۶] بکار رفت و سپس در مطالعات متعددی در قالب کاربردهای مختلف از قبیل کنترل پیش‌بین مدل متمرکز [۱۷]، کنترل پیش‌بین مدل نامتمرکز [۱۸]، کنترل پیش‌بین مدل سلسله‌مراتبی [۱۹]، کنترل پیش‌بین مدل تطبیقی [۳] و کنترل پیش‌بین مدل مقاوم [۲۰] مورد استفاده قرار گرفت. کنترل پیش‌بین مدل استاندارد برای محاسبه عمل کنترلی به مدل دقیق سیستم وابسته است و در نتیجه، در صورت وجود عدم قطعیت در پارامترهای مدل و دینامیک مدل نشده، عملکرد کنترلی می‌تواند تا حد زیادی تنزل یابد [۲۱]. علاوه بر آن، بار محاسباتی بالای حاصل از ساختار پیچیده روش

¹ Load frequency control

² Proportional-Integral-Derivative

³ Model predictive control

برنامه‌ریزی مربعی، سبب تاخیر در زمان پاسخ می‌شود و می‌تواند کاربرد این طرح در سیستم‌های قدرت واقعی را با مشکل مواجه کند.

در [۲۲]، یک روش کنترل پیش‌بین مدل مقاوم برای سیستم‌های قدرت چند ناحیه‌ای در حضور سیستم‌های ذخیره انرژی ترکیبی برای مساله کنترل بار-فرکانس پیشنهاد شده است. روش پیشنهادی بر مبنای یک قید برای مقاوم‌سازی قرار دارد که به‌طور موثری اغتشاشات نامعلوم سیستم را با فشرده‌سازی یک مجموعه تغییرناپذیر اغتشاش مدیریت می‌کند. در [۲۳]، یک روش کنترل پیش‌بین مدل نامتمرکز برای مساله کنترل بار-فرکانس سیستم‌های قدرت چند ناحیه‌ای با ژنراتورهای حرارتی و سیستم ذخیره انرژی پیشنهاد شده است و سیگنال بهینه کنترلی با روش برنامه‌ریزی مربعی خطی به دست آمده است. در [۲۴]، یک طرح کنترل بار-فرکانس با ترکیب واحد پیش‌بینی توان باد و کنترل پیش‌بین مدل طراحی شده است که به‌طور انعطاف‌پذیری به تغییرات تصادفی توان باد پاسخ می‌دهد. به منظور کاهش بار محاسباتی از توابع لاگرم^۱ در فرمول‌بندی مساله پیش‌بین مدل استفاده شده است. در [۲۱]، به منظور غلبه بر عدم قطعیت‌های مربوط به تولید و پارامترهای الکترومکانیکی ناشی از نفوذ بالای انرژی تجدیدپذیر، یک طرح کنترل بار-فرکانس پیش‌بین مدل بر مبنای داده برای یک سیستم قدرت چند ناحیه‌ای در حضور منابع بادی پیشنهاد شده است؛ یک مدل با استفاده از داده‌های ورودی-خروجی بصورت بازگشتی یادگیری می‌شود و بجای مدل پیش‌بینی دقیق مورد استفاده قرار می‌گیرد. در یک کار مشابه، در [۳]، یک طرح کنترل بار-فرکانس پیش‌بین مدل تطبیقی برای سیستم‌های قدرت چند ناحیه‌ای با لختی چرخشی نامعلوم و متغیر با زمان طراحی شده است؛ در ابتدا یک مدل اولیه با استفاده از داده‌های ورودی-خروجی به‌طور آفلاین به دست می‌آید و سپس با استفاده از روش تخمین افق متحرک در هر پله زمانی به‌طور آنلاین به‌روزرسانی می‌شود. در [۲۵]، یک طرح کنترل پیش‌بین مدل اقتصادی نامتمرکز برای مساله کنترل بار-فرکانس سیستم‌های قدرت چند ناحیه‌ای با توربین‌های بادی معرفی شده است؛ سیستم قدرت به چندین زیرسیستم با ترویج دینامیکی تقسیم و طراحی کنترل بهینه برای هر ناحیه با توجه به اطلاعات محلی همان ناحیه انجام شده است. در طرح مذکور، مسیرهای حالت آتی زیرسیستم‌های مجاور با استفاده از اطلاعات انتقالی بین نواحی تخمین زده می‌شود. با وجود نتایج قابل قبول کنترل‌های پیش‌بین مدل مذکور در تنظیم فرکانس در سیستم‌های قدرت بهم‌پیوسته چند ناحیه‌ای، تمام محدودیت‌های کنترل پیش‌بین مدل استاندارد از قبیل بار محاسباتی بالا (در نتیجه استفاده از مدل کامل برای پیش‌بینی مسیرهای آتی سیستم و لزوم تعداد نسبتاً بالای متغیرهای بهینه‌سازی برای حفظ عملکرد)، مدل ثابت پیش‌بینی و مشخصه مقاومت کنترلی ضعیف در حضور عدم قطعیت و تغییرات پارامتری در یک کار مورد بررسی قرار نگرفته است.

با توجه به چالش‌های مذکور در طراحی کنترل پیش‌بین مدل، در این مقاله یک روش کنترل پیش‌بین مدل تطبیقی^۲ (AMPC) برای بهبود مقاومت کنترلی در حضور تغییرات در پارامترهای سیستم قدرت و کاهش بار محاسباتی پیشنهاد شده است. موارد برجسته پژوهش در ادامه فهرست شده است:

- ساختار کنترلی شامل یک روش تخمین پارامتر بر مبنای تخمینگر مدل چند-جمله‌ای بازگشتی^۳ (RPME) و مدل خود رگرسیون با ورودی خارجی^۴ (ARX) برای ایجاد یک مدل کاهش مرتبه یافته از سیستم، فیلتر کالمن خطی تغییرپذیر با

¹ Laguerre functions

² Adaptive model predictive control

³ Recursive polynomial model estimator

⁴ Auto-regressive model with exogenous input

زمان^۱ (LTVKF) برای تخمین متغیرهای حالت سیستم و یک الگوریتم بهینه‌سازی با استفاده از توابع متعامد نرمال برای محاسبه عمل کنترلی با لحاظ قیود روی ورودی و خروجی سیستم می‌شود. روش کنترل پیشنهادی تنها بر مبنای طراحی یک کنترل‌کننده واحد است که به بار محاسباتی و فضای حافظه کمتری نیاز دارد و در عین حال، مشخصه مقاومت در برابر تغییرات در پارامترهای سیستم را بهبود می‌بخشد.

- بر خلاف کارهای انجام شده در [۲۲-۲۵]، یک روش شناسایی آنلاین داده‌محور در ساختار کنترلی گنجانده شده است؛ پارامترهای سیستم در هر گام کنترلی توسط واحد شناسایی به‌روزرسانی می‌شود تا با تغییرات دینامیکی در سیستم تطبیق یابند. بدین ترتیب، مقاومت طرح نسبت به تغییرات پارامتری به‌طور قابل ملاحظه‌ای بهبود می‌یابد. علاوه بر آن، استفاده از فیلتر کالمن متغیر با زمان می‌تواند در کاهش اثر نویز اندازه‌گیری روی عملکرد سیستم موثر واقع شود.
- بر خلاف روش کنترل مدل پیش‌بین تطبیقی در [۳ و ۲۶]، که یک مدل با مرتبه کامل از سیستم اصلی را شناسایی می‌کند، مدل شناسایی شده در کار پیشنهادی کاهش مرتبه یافته است. مدل مذکور می‌تواند مشخصه‌های اصلی رفتار سیستم را با دقت مطلوب بازتولید کند. در نتیجه، بار محاسباتی تا حد زیادی کاهش پیدا می‌کند و کیفیت مدل تخمینی نیز حفظ می‌شود. همچنین، در مقایسه با کارهای مراجع [۲۷-۲۸]، دنباله کنترلی آتی با استفاده از توابع متعامد نرمال لاگر تخمین زده می‌شود تا با کاهش تعداد متغیرهای بهینه‌سازی، به کاهش بیشتر بار محاسباتی کمک کند.
- کارایی روش کنترل پیش‌بین مدل تطبیقی پیشنهادی در دفع اغتشاش بار یک سیستم قدرت بهم‌پیوسته سه ناحیه‌ای با منابع حرارتی در حضور قید نرخ تولید^۲ (GRC) با استفاده از دو سناریوی پارامترهای نامی و با پارامترهای دارای عدم قطعیت بررسی شد و با عملکرد روش کنترل پیش‌بین مدل استاندارد MPC و کنترل PID مورد مقایسه قرار گرفت.

۲- مدل سیستم قدرت

دینامیک ناحیه i یک سیستم قدرت بهم‌پیوسته در شکل ۱ را می‌توان در حالت کلی با رابطه‌های (۱)-(۴) توصیف کرد [۲۹]:

$$\Delta \dot{f}_i = \frac{1}{2H_i} (\Delta P_{m,i} - \Delta P_{L,i} - \Delta P_{tie,i} - D_i \Delta f_i) \quad (1)$$

$$\Delta \dot{P}_{m,i} = \frac{1}{T_{t,i}} (\Delta P_{g,i} - \Delta P_{m,i}) \quad (2)$$

$$\Delta \dot{P}_{g,i} = \frac{1}{T_{g,i}} \left(\Delta P_{C,i} - \Delta P_{g,i} - \frac{1}{R_i} \Delta f_i \right) \quad (3)$$

$$\Delta \dot{P}_{tie,i} = 2\pi \sum_{j=1, j \neq i}^M T_{ij} (\Delta f_i - \Delta f_j) \quad (4)$$

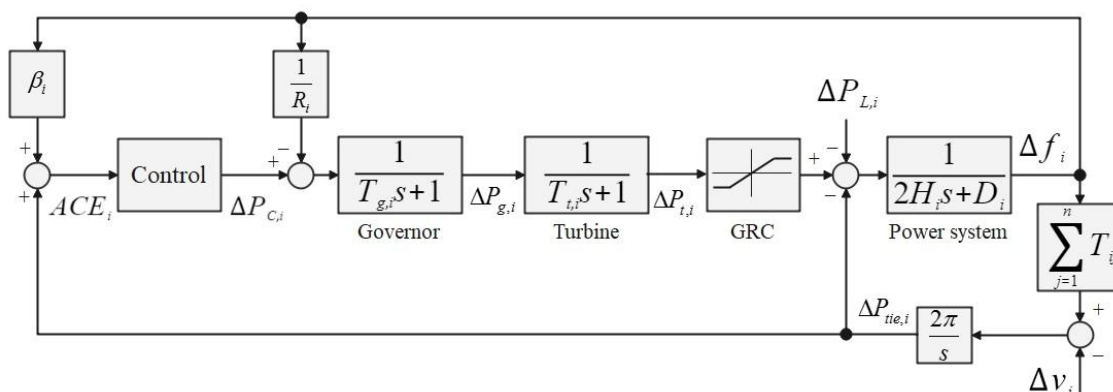
¹ Linear time-varying Kalman filter

² Generation rate control

که در آن Δf_i انحراف فرکانس در ناحیه i ، $\Delta P_{m,i}$ و $\Delta P_{g,i}$ به ترتیب توان مکانیکی توربین و خروجی گاورنر ژنراتور معادل و $\Delta P_{L,i}$ ، $\Delta P_{C,i}$ و $\Delta P_{tie,i}$ به ترتیب اغتشاش بار، سیگنال کنترل فرکانس بار و انحراف توان خط ارتباطی است. R_i ، D_i ، H_i ، $T_{t,i}$ و $T_{g,i}$ به ترتیب لختی سنکرون، ضریب میرایی، دروپ سرعت، ثابت زمانی ژنراتور و ثابت زمانی توربین ژنراتور معادل ناحیه i و ثابت سنکرون خط ارتباطی ناحیه i با ناحیه j است. خطای کنترل ناحیه i (ACE_i) طبق رابطه (۵) تعریف می‌شود:

$$ACE_i = \Delta P_{tie,i} + \beta_i \Delta f_i \quad (5)$$

که در آن β_i ضریب بایاس است. تضمین همگرایی ACE_i به صفر در حالت ماندگار و پاسخ دینامیکی مطلوب برای فرکانس از اهداف اصلی کنترل فرکانس بار است.



شکل ۱: مدل یک ناحیه در سیستم قدرت چند ناحیه‌ای.

معادلات (۱)–(۵) را می‌توان در قالب معادلات فضای حالت زمان-پیوسته نوشت [۳۰]:

$$\begin{cases} \dot{x}_i(t) = A_i x_i(t) + B_i u_i(t) + F_i d_i(t) \\ y_i(t) = C_i x_i(t) + D_i u_i(t) \\ |\dot{u}_i(t)| \leq r_{\max} \end{cases} \quad (6)$$

که در آن $x_i(t) = [\Delta P_{g,i}, \Delta P_{m,i}, \Delta f_i, \Delta P_{tie,i}]^T$ بردار متغیرهای حالت، $u_i(t) = \Delta P_{C,i}$ ورودی کنترل، $d_i = [\Delta P_{L,i}, \Delta v_i]^T$ بردار ورودی اغتشاش، $\Delta v_i = \sum_{j=1, j \neq i}^n T_{ij} \Delta f_j$ ، $y_i(t) = ACE_i$ خروجی سیستم و r_{\max} قید نرخ تولید است. ماتریس‌های مدل کامل ناحیه i به کمک معادلات (۱)–(۵) به دست می‌آیند [۳۰]:

$$A_i = \begin{bmatrix} -1/T_{g,i} & 0 & -1/R_i T_{g,i} & 0 \\ 1/T_{t,i} & -1/T_{t,i} & 0 & 0 \\ 0 & 1/2H_i & -D_i/2H_i & -1/2H_i \\ 0 & 0 & 2\pi \sum_{j=1, j \neq i}^n T_{ij} & 0 \end{bmatrix}, F_i = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -1/(2H_i) & 0 \\ 0 & -2\pi \end{bmatrix} \quad (7)$$

$$B_i = [1/T_{g,i} \ 0 \ 0 \ 0]^T, C_i = [0 \ 0 \ \beta_i \ 1]$$

¹ Area control error

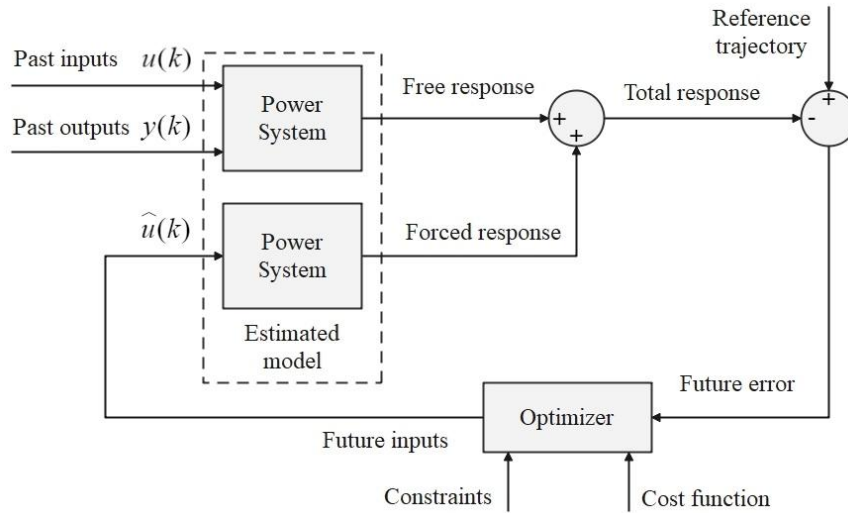
۳- طراحی سیستم کنترل

یک روش کنترل پیش‌بین مدل تطبیقی برای مساله کنترل بار-فرکانس سیستم‌های قدرت چند ناحیه‌ای در حضور تغییرات در پارامترهای سیستم و اغتشاش بار طراحی شده است. مدل خطی با کاهش مرتبه شناسایی شده سیستم قدرت در قانون کنترل پیش‌بین مدل تطبیقی مورد استفاده قرار می‌گیرد تا فرکانس هر ناحیه و توان تبدیلی در خط ارتباطی را، با تغییر در میزان بار مرجع تمام واحدهای تولیدی در هر ناحیه، تنظیم کند.

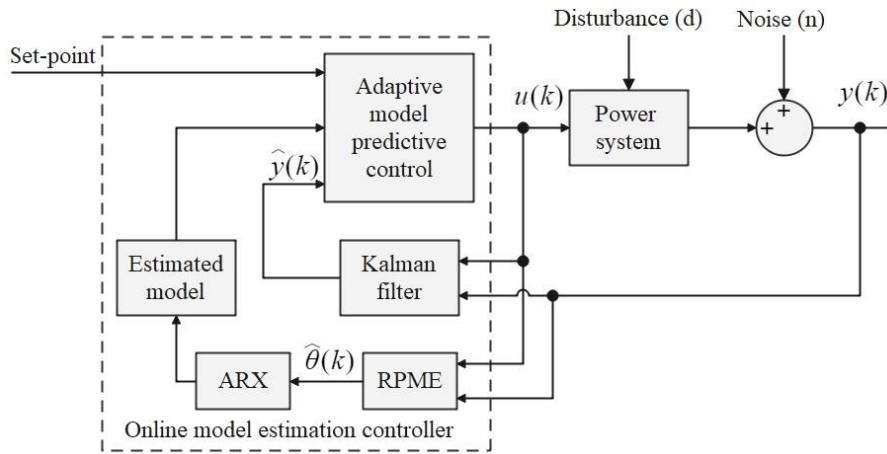
۳-۱- کنترل پیش‌بین مدل تطبیقی

کنترل پیش‌بین مدل یک ابزار قوی برای حل مسایل کنترلی با وجود قیود و چرخه‌های کنترلی طولانی است. این روش کنترلی به‌طور موفق روی کاربردهای صنعتی مختلفی اجرا شده است. اگرچه این روش کنترلی برای کاربردهای مختلف به منظور دستیابی به عمل کنترلی بهینه مناسب است، سیستم‌های فیزیکی واقعی دارای مشخصه‌های غیرخطی قابل ملاحظه‌ای هستند. علاوه بر آن، مشخصه دینامیکی این سیستم‌ها ممکن است به‌طور آهسته یا سریع در زمان تغییر کند و با توجه به اینکه ساختار کنترل‌کننده پیش‌بین مدل استاندارد تطبیقی نیست، مقاومت آن در برابر تغییرات پارامترهای سیستم محدود است. با استفاده از این کنترل‌کننده در حضور عدم قطعیت در مدل، مسیر پیش‌بینی شده سیستم خیلی دقیق نخواهد بود و در نتیجه، عملکرد کنترلی تنزل می‌یابد.

یک رهیافت متداول برای رفع این عیب، اجرای کنترل پیش‌بین مدل تطبیقی است که امکان تخمین پارامترهای سیستم و به‌روزرسانی آنالین مدل پیش‌بینی را فراهم می‌آورد. مسیر آتی سیستم در کنترل پیش‌بین مدل تطبیقی در ابتدا با یک مدل خطی تغییرناپذیر با زمان در فضای حالت به دست می‌آید؛ اما مدل مذکور در طول زمان می‌تواند با توجه به شرایط سیستم به‌روزرسانی شود. ایده کلی کنترل پیش‌بین مدل با مدل تخمینی سیستم در شکل ۲ نشان داده شده است. عملکرد کنترل پیش‌بین مدل به افق پیش‌بینی و قیود خروجی بستگی دارد. در هر بازه کنترلی، مسیر آتی سیستم با استفاده از یک مدل پیش‌بینی و متغیرهای حالت فعلی سیستم (بعنوان شرایط اولیه) به دست می‌آید و یک مساله بهینه‌سازی زمان گسسته با ابعاد محدود به منظور حصول دنباله کنترلی آتی حل می‌شود [۳۱]. ساختار کنترل‌کننده پیش‌بین مدل تطبیقی در ترکیب با مدل سیستم قدرت، شامل فیلتر کالمن تغییرپذیر با زمان LTVKF، مدل خود رگرسیون با ورودی خارجی ARX و تخمینگر مدل چندجمله‌ای بازگشتی RPME در شکل ۳ تصویر شده است. تخمینگر مدل چندجمله‌ای بازگشتی به منظور تخمین پارامترهای سیستم در زمان حقیقی برای شناسایی مدل چندجمله‌ای ARX و استفاده از آن در به‌روزرسانی مدل پیش‌بینی کنترل پیش‌بین مدل بکار می‌رود.



شکل ۲. ایده کلی کنترل پیش‌بین مدل.



شکل ۳. ساختار کنترل پیش‌بین مدل تطبیقی پیشنهادی.

توصیف مدل خطی زمان گسسته هر ناحیه سیستم قدرت در فضای حالت را در نظر می‌گیریم که با استفاده از مدل ARX (در بخش ۳-۲) شناسایی شده است:

$$x(k+1) = A_{id}^k x(k) + B_{id}^k u(k) + L(k)d(k) \quad (8)$$

$$y(k) = C_{id}^k x(k) + n(k) \quad (9)$$

که در آنها $A_{id}^k \in \mathbb{R}^{2 \times 2}$ ماتریس حالت، $B_{id}^k \in \mathbb{R}^{2 \times 1}$ ماتریس ورودی و $C_{id}^k \in \mathbb{R}^{1 \times 2}$ ماتریس سنجش است. همچنین، $x(k)$ ، $y(k)$ ، $u(k)$ و $L(k)$ به ترتیب تخمین بردار حالت سیستم، خروجی سیستم، متغیر کنترلی، پله زمانی و بهره کالمن است. علاوه بر آن، اغتشاش سیستم $d(k)$ با ماتریس کواریانس $Q(k)$ و نویز اندازه‌گیری $n(k)$ با ماتریس کواریانس $R(k)$ از نوع نویز سفید گوسی با مقدار متوسط صفر و کواریانس متقابل $N(k)$ هستند:

$$Q(k) = E[d(k)d(k)^T], R(k) = E[n(k)n(k)^T], N(k) = E[d(k)n(k)^T], \quad (10)$$

که در آن E امید ریاضی متغیر تصادفی است. بایستی خاطرنشان کرد که اغتشاش سیستم $d(k)$ و نویز اندازه‌گیری $n(k)$ دارای همبستگی هستند. اغتشاش سیستم در اینجا تغییرات بار و اثر نواحی مجاور روی هر ناحیه و نویز اندازه‌گیری همان نویز حسگرها است. فرض بر آن است که ماتریس‌های کواریانس $Q(k)$ ، $R(k)$ و $N(k)$ قطری هستند و برای تنظیم بهره فیلتر کالمن، نسبت به زمان تغییر می‌کنند. می‌توان $d(k)$ و $n(k)$ را در قالب معادلات فضای حالت توصیف کرد:

$$x_d(k+1) = A_d^k x_d(k) + B_d^k u_d(k) \quad (11)$$

$$d(k) = C_d^k x_d(k) \quad (12)$$

$$x_n(k+1) = A_n^k x_n(k) + B_n^k u_n(k) \quad (13)$$

$$n(k) = C_n^k x_n(k) \quad (14)$$

که در آنها A_d^k ، B_d^k ، C_d^k ، A_n^k ، B_n^k ، C_n^k ماتریس‌های مربوط به اغتشاش سیستم و نویز اندازه‌گیری هستند که در زمان حقیقی به‌روزرسانی می‌شوند و $u_d(k)$ و $u_n(k)$ به ترتیب اغتشاش ورودی و نویز ورودی است. این ورودی‌های مدل فضای حالت از نوع نویز سفید گوسی با متوسط صفر و واریانس واحد هستند. با افزودن تخمین‌های نویز اندازه‌گیری $n(k)$ و اغتشاش سیستم $d(k)$ به مدل شناسایی شده سیستم، یک مدل تعمیم‌یافته طبق رابطه‌های (۱۵) و (۱۶) به دست می‌آید:

$$x_c(k+1|k) = A_c^k x_c(k) + B_c^k u_c(k) \quad (15)$$

$$y(k) = C_c^k x_c(k) \quad (16)$$

که در آنها $x_c = [x(k), x_d(k), x_n(k)]^T$ و $u_c = [u(k), u_d(k), u_n(k)]^T$ به ترتیب بردار حالت و ورودی مدل تعمیم‌یافته هستند و ماتریس‌های سیستم با رابطه (۱۷) به دست می‌آیند:

$$A_c^k = \begin{bmatrix} A_{id}^k & B_{id,d}^k & 0 \\ 0 & A_d^k & 0 \\ 0 & 0 & A_n^k \end{bmatrix}, B_c^k = \begin{bmatrix} B_{id}^k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & B_n^k \end{bmatrix}, C_c^k = [C_{id}^k \quad 0 \quad C_n^k] \quad (17)$$

که در آن $B_{id,d}^k$ ماتریس اغتشاش ورودی است. بنابراین، تخمین بردار حالت و خروجی سیستم در پله زمانی فعلی با استفاده از بردار سنجش در پله زمانی k با رابطه‌های (۱۸) و (۱۹) توصیف می‌شود:

$$x_c(k|k) = x_c(k|k-1) + M(k)(y(k) - C_c^k x_c(k|k-1)) \quad (18)$$

$$\hat{y}(k|k) = C_c^k x_c(k|k) \quad (19)$$

که در آنها $\hat{y}(k|k)$ تخمین خروجی سیستم در پله زمانی فعلی است. تخمین حالت مدل و خروجی متناظر برای پله کنترلی $k+1$ با رابطه‌های (۲۰)–(۲۲) به دست می‌آیند:

$$x_c(k+1|k) = A_c^k x_c(k|k) + B_c^k u_c(k) + L(k)e(k) \quad (20)$$

$$\hat{y}(k|k) = C_c^k x_c(k+1|k) \quad (21)$$

$$e(k) = y(k) - C_c^k x_c(k+1|k) \quad (22)$$

بایستی خاطر نشان کرد که عبارت $L(k)e(k)$ در رابطه (۲۰) برای حصول قانون کنترل صحیح در فرمول‌بندی بهینه‌سازی کنترل پیش‌بین مدل نقش مهمی دارد [۳۲]. در هر بازه کنترلی، بهره‌های کالمن $L(k)$ و $M(k)$ به منظور تطبیق مدل تخمینی با سیستم واقعی تنظیم می‌شوند. بهره‌های مذکور با استفاده از تخمین قبلی ماتریس کواریانس حالت و حل معادله ریکاتی جبری طبق رابطه‌های (۲۳)–(۲۵) تعیین می‌شوند [۳۳]:

$$L(k) = \left(A_c^k P(k|k-1) C_c^{kT} + N(k) \right) \left(C_c^k P(k|k-1) C_c^{kT} + R(k) \right)^{-1} \quad (23)$$

$$M(k) = P(k|k-1) C_c^{kT} \left(C_c^k P(k|k-1) C_c^{kT} + R(k) \right)^{-1} \quad (24)$$

$$P(k+1|k) = A_c^k P(k|k-1) A_c^{kT} - \left(A_c^k P(k|k-1) C_c^{kT} + N(k) \right) + Q(k) \quad (25)$$

$$P(k) = \left(I - M(k) C_c^k \right) P(k|k-1) \quad (26)$$

که در آنها $P(k|k-1)$ ماتریس کواریانس خطای تخمین حالت در پله زمانی $k+1$ است که با استفاده از اطلاعات موجود در پله زمانی k به دست می‌آید. رابطه (۲۳)، ماتریس بهره کالمن بهینه را برای یک سیستم با اغتشاش و نویز اندازه‌گیری دارای همبستگی محاسبه می‌کند. هنگامی که تخمین پارامترهای مدل در زمان حقیقی با استفاده از تخمینگر مدل چندجمله‌ای بازگشتی به دست آید و با استفاده از مدل ARX به قالب فضای حالت تبدیل شوند، ماتریس‌های کواریانس اغتشاش $d(k)$ و نویز $n(k)$ ، یعنی ماتریس‌های $R(k)$ ، $Q(k)$ و $N(k)$ ، در فضای حالت نیز با زمان تغییر می‌کنند. بنابراین، اغتشاش سیستم و نویز اندازه‌گیری در معادلات (۱۱)–(۱۴) را می‌توان با استفاده از یک فیلتر کالمن تخمین زد:

$$\hat{x}_d(k|k) = \hat{x}_d(k|k-1) + M(k)(y - \hat{y}) \quad (27)$$

$$\hat{x}_n(k|k) = \hat{x}_n(k|k-1) + M(k)(y - \hat{y}) \quad (28)$$

و تخمین متغیرهای حالت مذکور برای بازه کنترلی بعدی، $k+1$ ، با استفاده از رابطه‌های (۲۹) و (۳۰) به دست می‌آید:

$$\hat{x}_n(k+1|k) = A_n \hat{x}_n(k|k) \quad (29)$$

$$\hat{x}_d(k+1|k) = A_d \hat{x}_d(k|k) \quad (30)$$

همانطور که قبلاً ذکر شد، کارایی کنترل پیش‌بین مدل تا حد زیادی به دقت پیش‌بینی خروجی سیستم بستگی دارد. با استفاده از فیلتر کالمن، می‌توان تضمین کرد که اطلاعات خروجی با نویز کم در اختیار کنترل‌کننده قرار گیرد. بدون استفاده از فیلتر کالمن، نویز اندازه‌گیری در مساله بهینه‌سازی وارد می‌شود که می‌تواند عملکرد کنترلی را کاهش دهد. علاوه بر آن، فیلتر کالمن می‌تواند تخمینی از متغیرهای حالت غیر قابل سنجش را فراهم کند.

۲-۳- شناسایی و تخمین آنلاین پارامترهای مدل

بار محاسباتی بالای شبیه‌سازی مدل دینامیکی غیرخطی و پیچیده سیستم قدرت سبب می‌شود تا استفاده از آن به‌طور مستقیم در چارچوب کنترل پیش‌بین مدل با دشواری همراه باشد. بنابراین، یک مدل ساده شده که مشخصه‌های کلی رفتار سیستم قدرت واقعی را بازسازی می‌کند به منظور پیش‌بینی رفتار آتی سیستم در کنترل پیش‌بین مدل مورد نیاز است. شناسایی سیستم شامل ایجاد مدل‌های دینامیکی بر مبنای داده مانند مدل ARX است، تا رفتار دینامیکی متغیرهای موردنظر را پیش‌بینی کند.

یک معادله تفاضلی خطی شاید ساده‌ترین فرم یک سیستم خطی گسسته در زمان باشد:

$$y(k) + a_1 y(k-1) + \dots + a_{n_a} y(k-n_a) = b_1 u(k-1) + \dots + b_{n_b} u(k-n_b) + e(k) \quad (31)$$

که در آن $a_i, i=1, \dots, n_a$ و $b_i, i=1, \dots, n_b$ پارامترهای مدل هستند که بایستی تخمین زده شوند و n_a, n_b مرتبه‌های مدل هستند.

توصیف کلی مدل چند جمله‌ای ARX در رابطه (۳۱) طبق رابطه (۳۲) قابل بیان است:

$$A(z)y(k) = B(z)u(k) + e(k) \quad (32)$$

که در آن $u(k)$ و $y(k)$ به ترتیب ورودی و خروجی مدل سیستم هستند؛ $e(k)$ نویز سفید با متوسط و واریانس صفر است و $A(z)$ و $B(z)$ چند جمله‌هایی بر حسب عملگر شیفت زمانی رو به عقب هستند:

$$A(z) = 1 + a_1(k)z^{-1} + a_2(k)z^{-2} + \dots + a_{n_a}(k)z^{-n_a} \quad (33)$$

$$B(z) = b_1(k)z^{-1} + b_2(k)z^{-2} + \dots + b_{n_b}(k)z^{-n_b} \quad (34)$$

$$z^{-1}y(k) = y(k - \Delta k) = y(k-1) \quad (35)$$

که در آنها z^{-1} عملگر شیفت زمانی رو به عقب و Δk بازه نمونه‌برداری است. فرض کنید که مرتبه‌های n_a و n_b معلوم هستند و مدل (۳۱) را می‌توان در هر زمان k در قالب شبه‌رگرسیون بیان کرد:

$$y(k) = \phi^T(k)\theta(k) + e(k) \quad (36)$$

که در آن $\theta(k) \in \mathbb{R}^n$ بردار پارامترهایی هستند که بایستی تخمین زده شوند:

$$\theta(k) = [a_1(k), a_2(k), \dots, a_{n_a}(k), b_1(k), b_2(k), \dots, b_{n_b}(k)]^T \quad (37)$$

و $\phi(k) \in \mathbb{R}^n$ بردار گرادیان متغیرهای ورودی و خروجی است:

$$\phi(k) = [-y(k-1), -y(k-2), \dots, -y(k-n_a), u(k-1), u(k-2), \dots, u(k-n_b)]^T \quad (38)$$

با توجه به بردار پارامترها در رابطه (۳۷)، یک مدل تک-ورودی تک-خروجی مرتبه دوم برای هر ناحیه از سیستم قدرت در نظر گرفته شده است که طبق رابطه (۳۹) بیان می‌شود:

$$y(k+1) = -a_1 y(k-1) - a_2 y(k-2) + b_1 u(k-1) + b_2 u(k-2) + e(k) \quad (39)$$

که در آن a_1, a_2, b_1, b_2 پارامترهای متغیر با زمان مدل هستند.

تحقق فضای حالت یک مدل ARX با پارامترهای ثابت در رابطه (۳۹) به یک قالب مدل فضای حالت برای فیلتر کالمن حالت ماندگار در رابطه‌های (۸) و (۹) منجر می‌شود [۳۴].

به منظور تخمین پارامترهای سیستم θ و ماتریس کواریانس مربوطه $Q(k)$ در زمان حقیقی، رهیافت تخمینگر مدل چندجمله‌ای بازگشتی با الگوریتم تطبیق ضریب فراموشی مورد استفاده قرار می‌گیرد [۳۴]. این رهیافت برای سازگاری با مدل ARX زمان گسسته شناخته شده است [۳۵]. بردار پارامترهای تخمینی $\hat{\theta}(k)$ طبق رابطه (۴۰) توصیف می‌شود:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + K(k)(y(k) - \hat{y}(k)) \quad (40)$$

که در آن $K(k)$ بهره‌ای است که روی به‌روزرسانی بردار پارامترها اثر می‌گذارد:

$$K(k) = Q(k)\phi(k) \quad (41)$$

که در آن $\phi(k)$ بردار گرادیان $\hat{\theta}(k)$ است و $Q(k)$ با رابطه (۴۲) به دست می‌آید:

$$Q(k) = \frac{P_{ide}(k-1)}{\gamma + \phi^T(k)P_{ide}(k-1)\phi(k)} \quad (42)$$

طبق معادلات رگرسیون خطی، تخمین خروجی با استفاده از تخمین بردار پارامترها $\hat{\theta}(k)$ حاصل می‌شود:

$$\hat{y}(k) = \phi^T(k)\hat{\theta}(k-1) \quad (43)$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{P_{ide}(k-1)\phi(k)}{\gamma + \phi^T(k)P_{ide}(k-1)\phi(k)} (y(k) - \phi^T(k)\hat{\theta}(k-1)) \quad (44)$$

که در آن $P_{ide}(k-1)$ ماتریس کواریانس خطای شناسایی پارامترها و γ ضریب فراموشی است.

برای اجتناب از بار محاسباتی در نتیجه محاسبه معکوس ماتریس کواریانس، $P_{ide}(k)$ با استفاده از رابطه (۴۵) به دست می‌آید:

$$P_{ide}(k) = \frac{1}{\gamma} \left[P_{ide}(k-1) - \frac{P_{ide}(k-1)\phi(k)\phi^T(k)P_{ide}(k-1)}{\gamma + \phi^T(k)P_{ide}(k-1)\phi(k)} \right] \quad (45)$$

روش تخمینگر مدل چندجمله‌ای بازگشتی با ضریب فراموشی واحد ($\gamma=1$) می‌تواند دقت تخمین را کاهش دهد. برای سیستم‌های غیرخطی متغیر با زمان، تخمینگر مدل چندجمله‌ای بازگشتی با ضریب فراموشی نمایی ($\gamma < 1$) به نتایج قابل قبول منجر خواهد شد [۳۶]. در مراحل اولیه محاسبات، بهتر است که ضریب فراموشی در یک مقدار کوچک تنظیم شود. این کار کمک می‌کند تا اثر خطای بزرگ بین مقادیر تخمینی قبلی و واقعی روی دقت تخمین پارامتر کاهش یابد. هنگامی که خطای تخمین پارامتر در ادامه روند محاسبات کوچک‌تر می‌شود، بهتر است تا ضریب فراموشی را افزایش دهیم و در مقادیر نزدیک به واحد تنظیم کنیم. تخمین بردار پارامتر $\hat{\theta}(k)$ برای هر بازه کنترلی تعیین می‌شود و سپس، چندجمله‌ای‌های $A(z)$ و $B(z)$ در قالب توصیف فضای حالت به ماتریس‌های A_{id}^k ، B_{id}^k و C_{id}^k تبدیل می‌شوند که در معادلات (۸) و (۹) مورد استفاده قرار گرفته‌اند. الگوریتم تخمینگر مدل چندجمله‌ای بازگشتی در صورتی در تخمین بردار پارامتر $\hat{\theta}(k)$ خوب عمل می‌کند که متغیرهای ورودی و خروجی رویت‌پذیر باشند. اگر پارامترهای سیستم به کندی نسبت به زمان تغییر کنند، تاثیر داده‌های قبلی روی بردار پارامتر می‌تواند با افزایش سطح اطمینان بردار داده‌های ورودی-خروجی نزدیک به زمان فعلی، کاهش یابد [۳۷]. زمانی که داده‌های نمونه ورودی-خروجی کافی جمع‌آوری شد، مساله حداقل مربعات برای تعیین پارامترهای مدل چندجمله‌ای ARX حل می‌شود. این کار با کمینه‌سازی تابع هدف زیر انجام می‌شود:

$$J_k(\hat{\theta}) = \frac{1}{2} \sum_{i=1}^k \gamma^{k-i} [y_i(k) - \phi^T(k)\hat{\theta}(k)]^2 \quad (46)$$

از آنجایی که $J_k(\hat{\theta})$ یک تابع مربعی اکیدا محدب است، $\theta(k+1) = \arg \min_{\hat{\theta} \in \mathbb{R}^{n \times n}} J_k(\hat{\theta})$ تنها پاسخ مساله بهینه‌سازی است. مراحل تخمین پارامترهای اولیه مدل با استفاده از معادلات تخمینگر مدل چندجمله‌ای بازگشتی در ادامه فهرست شده است. مراحل تخمین پارامترهای اولیه مدل با استفاده از معادلات تخمینگر مدل چندجمله‌ای بازگشتی در ادامه فهرست شده است:

۱. مقداردهی اولیه: $k = 1$ ، بیشینه طول داده‌ها $L_{data} = 1200$ ، $\gamma < 1$ ، $\phi(0) \in \mathbb{R}^{n \times n}$ ، $\theta(0) \in \mathbb{R}^{n \times n}$ ، $P(0) = p_0 I_{n \times n}$ ، $p_0 = 10^6$.
۲. ایجاد بردار داده‌های سنجش ورودی-خروجی $\phi(k)$ با استفاده از رابطه (۳۸).
۳. محاسبه بردار بهره $K(k)$ و ماتریس کواریانس $P(k-1)$ و $P(k)$ با استفاده از معادلات (۳۸)، (۴۰)–(۴۲) و (۴۵).
۴. به‌روزرسانی بردار تخمین پارامترها $\hat{\theta}(k)$ با استفاده از رابطه (۴۴).
۵. اگر $k < L_{data}$ ، افزایش k به اندازه یک واحد و رفتن به مرحله ۲، در غیر اینصورت، خاتمه اجرای الگوریتم.

۳-۳- فرمول‌بندی مساله بهینه‌سازی کنترل پیش‌بین مدل

بهینه‌سازی یکی از بخش‌های اصلی کنترل‌کننده کنترل پیش‌بین مدل است که در ضمن آن بهترین توالی عمل‌های کنترلی آتی با بهینه‌سازی تابع هدف به دست می‌آید. یک الگوریتم کنترل پیش‌بین مدل بر مبنای مدل خطی سیستم به‌طور معمول از تابع هدف برای کمینه‌سازی خطای بین خروجی سیستم و خروجی مطلوب بهره می‌گیرد. تابع هزینه در معرض قیود مختلف تساوی و ناتساوی قرار دارد و برای یافتن بهترین توالی کنترلی بهینه‌سازی می‌شود. این بهینه‌سازی به‌طور تکراری در هر بازه کنترلی انجام می‌شود. تابع هزینه مربعی شامل جمع وزنی خطاهای پیش‌بینی و مقادیر ورودی کنترلی آتی می‌شود:

$$\Phi = \sum_{j=1}^{H_p} R_j (y_r(k) - \hat{y}(k+j|k))^2 + \sum_{j=1}^{H_c} S_j (\Delta u(k+j))^2 \quad (47)$$

که در آن $R_j = \text{diag}[r_j(k), \dots, r_j(H_p-1)]$ و $S_j = \text{diag}[s_j(k), \dots, s_j(H_c-1)]$ ضرایب تنظیم وزنی، H_p افق پیش‌بینی، H_c افق کنترلی، \hat{y} و y_r به ترتیب مقدار پیش‌بینی و مقدار مرجع متغیر خروجی، u متغیر کنترلی و $\Delta u(k+j) = u(k+j|k) - u(k+j-1|k)$ نمو ورودی در طول یک بازه زمانی است. خروجی سیستم توسط حدود بیشینه و کمینه محدود می‌شوند:

$$y_{\min} \leq \hat{y}(k+j) \leq y_{\max}, j = 1, \dots, H_p \quad (48)$$

علاوه بر آن، ورودی‌های کنترلی نیز دارای قیدهایی از قبیل محدودیت روی اندازه و محدودیت روی نرخ تغییر ورودی هستند:

$$\begin{cases} u_{\min} \leq u(k+j) \leq u_{\max} \\ \Delta u_{\min} \leq \Delta u(k+j) \leq \Delta u_{\max} \end{cases}, j = 1, \dots, H_c \quad (49)$$

به منظور کاهش بار محاسباتی، می‌توان نمو ورودی کنترلی را با استفاده از توابع لاگر تخمین زد:

$$\Delta u(k+j) = L(j)^T \eta = [L_1(j), \dots, L_N(j)] [\eta_1, \dots, \eta_N]^T \quad (50)$$

که در آن N مرتبه شبکه لاگر و $L_i(j)$, $i=1, \dots, N$ توابع پایه لاگر زمان گسسته است که با استفاده از یک معادله تفاضلی به دست می‌آید:

$$L(k+1) = A_l L(k), L(0)^T = \sqrt{\beta} [1, -\alpha, \alpha^2, -\alpha^3, \dots, (-1)^{N-1} \alpha^{N-1}], \beta = 1 - \alpha^2, \quad (51)$$

$$A_l = \begin{bmatrix} \alpha & 0 & \dots & 0 \\ \beta & \alpha & \dots & 0 \\ -\alpha\beta & \beta & \dots & 0 \\ \alpha^2\beta & -\alpha\beta & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ (-\alpha)^{N-2}\beta & (-\alpha)^{N-3}\beta & \dots & \alpha \end{bmatrix}$$

که در آن $L(0)$ مقدار اولیه و $0 \leq \alpha < 1$ ضریب مقیاس زمانی شبکه لاگر است.

به این ترتیب، بهره کنترل پیش‌بین مدل K_{mpc} با استفاده از رابطه (52) در هر گام کنترلی محاسبه می‌شود:

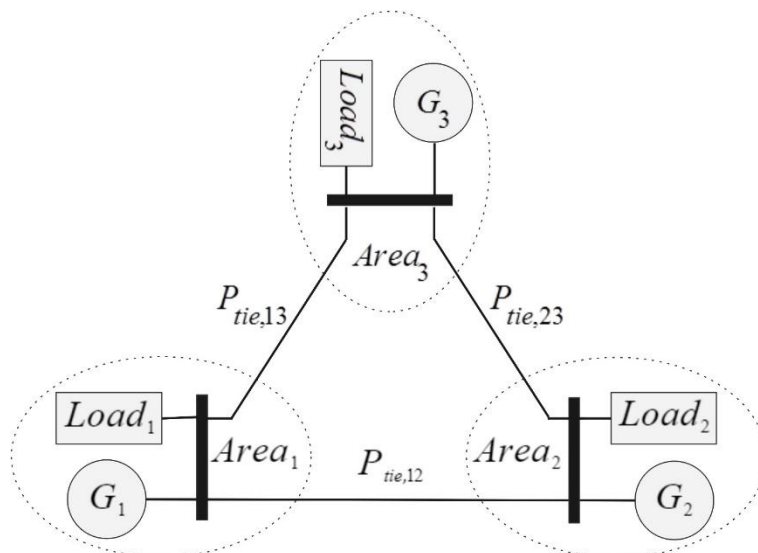
$$K_{mpc} = L(0)^T \Omega^{-1} \Psi, \Omega = \sum_{j=1}^{H_p} \varphi(j) Q \varphi(j)^T + R_L, \Psi = \sum_{j=1}^{H_p} \varphi(j) Q (A_c^k)^j, \quad (52)$$

$$\varphi(j) = \sum_{i=0}^j (A_c^k)^{j-i-1} B_c^k L(i)^T, Q = (C_c^k)^T C_c^k, R_L = r_w \text{diag}[1, \dots, 1]$$

۴- نتایج شبیه‌سازی

برای بررسی عملکرد کنترل‌کننده بار-فرکانس پیش‌بین مدل تطبیقی پیشنهادی، شبیه‌سازی‌های عددی در نرم‌افزار سیمولینک متلب انجام شد. ساختار قدرت سه ناحیه‌ای بهم‌پیوسته در شکل ۴ به عنوان سیستم اصلی مورد بررسی قرار گرفت. پارامترهای نامی مدل در جدول ۱ آورده شده است [۳۸]. مدل هر ناحیه در سیستم قدرت سه ناحیه‌ای مذکور با معادلات (۱)–(۵) توصیف می‌شود. ظرفیت کنترل فرکانس بار در مقدار 0.2 pu، گام نمونه‌برداری در مقدار $dt = 0.01$ s و با توجه به اینکه تمام واحدهای تولید توان از نوع ژنراتورهای حرارتی هستند، نرخ تغییر ورودی کنترل (قید نرخ تولید) در مقدار 0.05 pu/min تنظیم شده است. مقادیر اولیه پارامترهای مدل با استفاده از شناسایی مدل به‌طور آفلاین به دست آمد و با توجه به تغییرات پارامترهای سیستم در زمان حقیقی به‌روزرسانی خواهد شد. برای کنترل‌کننده پیش‌بین مدل تطبیقی پیشنهادی، افق پیش‌بینی $H_p = 20$ ، ضریب مقیاس زمانی $\alpha = 0.9$ ، مرتبه شبکه لاگر $N = 3$ و ضریب وزنی $r_w = 45$ است. این پارامترهای کنترلی برای هر سه ناحیه کنترلی در شبکه قدرت در شکل ۴ یکسان است. شبیه‌سازی‌ها با اغتشاش پله‌ای بار با مقدار $\Delta P_L = 0.02$ pu در ناحیه ۲ در زمان $t = 0$ s با دو سناریوی (۱) با پارامترهای نامی طبق جدول ۱ و (۲) پارامترهای دارای عدم قطعیت انجام شد. باید خاطر نشان شود که اغتشاش بار در ناحیه ۲ از طریق خطوط ارتباطی روی نواحی دیگر اثر می‌گذارد. با توجه به قید نرخ تولید یکسان، عملکرد کنترل پیشنهادی AMPC با کنترل پیش‌بین مدل استاندارد MPC و کنترل تناسبی-انتگرالی-مشتقی PID مقایسه شد. برای تمام موارد، شاخص زمان نشست به بازه زمانی بعد از اعمال اغتشاش بار اشاره دارد که پاسخ به محدوده ± 0.0001 حول صفر همگرا می‌شود. برای پارامترهای کنترل پیش‌بین مدل استاندارد، افق پیش‌بینی $H_p = 120$ و افق کنترلی $H_c = 10$ برای هر سه ناحیه کنترلی تنظیم شد. برای تنظیم پارامترهای افق کنترلی و افق پیش‌بینی در کنترل پیش‌بین مدل استاندارد، مواردی از قبیل عملکرد حالت نامی، مشخصه مقاوم و بار محاسباتی در نظر گرفته

شده است. پارامترهای کنترل PID به‌طور جداگانه و با توجه به دینامیک‌های هر ناحیه انتخاب شده است: $K_{P1} = 0.2726$ ، $K_{I3} = 0.6135$ ، $K_{P3} = 0.3276$ ، $K_{D2} = 0.25$ ، $K_{I2} = 0.7$ ، $K_{P2} = 0.8515$ ، $K_{D1} = 0.3213$ ، $K_{I1} = 0.4821$ و $N_1 = N_2 = N_3 = 100$ و $K_{D3} = 0.4852$.



شکل ۴: سیستم قدرت بهم‌پیوسته سه-ناحیه‌ای.

جدول ۱: پارامترهای سیستم قدرت سه-ناحیه‌ای [۳۸].

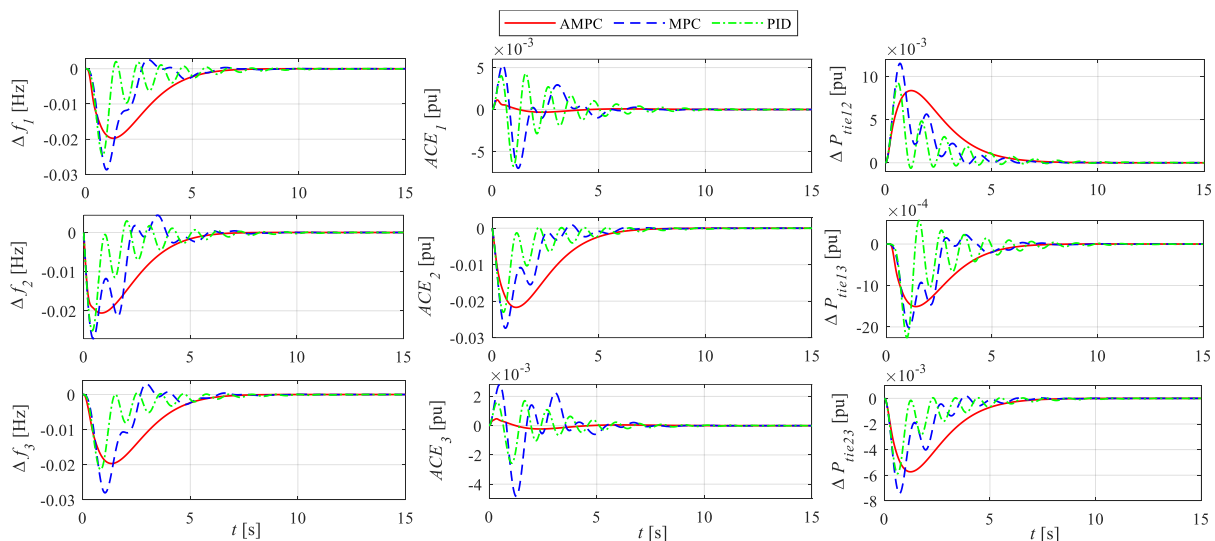
Area	d [pu/Hz]	$2H$ [pu s]	R [Hz/pu]	T_g [s]	T_r [s]	β [pu/Hz]	T_{ij} [pu/Hz]
1	0.015	0.1667	3.00	0.08	0.40	0.3483	$T_{12}=0.20, T_{13}=0.25$
2	0.016	0.2017	2.73	0.06	0.44	0.3827	$T_{21}=0.20, T_{23}=0.15$
3	0.015	0.1247	2.82	0.07	0.30	0.3692	$T_{31}=0.25, T_{32}=0.15$

۴-۱- سناریوی ۱: بررسی عملکرد با پارامترهای نامی

نتایج شبیه‌سازی در این سناریو در شکل ۵ نشان داده شده است که شامل انحراف فرکانس، انحراف توان در خط ارتباطی و خطای کنترل ناحیه (متغیر خروجی تحت کنترل) می‌شود. با توجه به این شکل، هموار بودن پاسخ سیستم تحت کنترل پیشنهادی نسبت به کنترل پیش‌بین مدل استاندارد و کنترل PID نشان دهنده میرایی بهتر آن است؛ پاسخ سیستم تحت کنترل پیش‌بین مدل و کنترل PID با نوسانات بزرگی به صفر همگرا می‌شوند. با توجه به اغتشاشات دائمی در سیستم قدرت ناشی از تغییرات بار، سرعت و میرایی پاسخ سیستم از نظر حفظ کیفیت توان و پایداری بسیار دارای اهمیت است. مقایسه عددی کنترل‌کننده‌های مورد بررسی از نظر زمان نشست (سرعت پاسخ حالت ماندگار) و بیشینه انحراف متغیرهای مذکور در جدول ۲ و نمودار میله‌ای آن در شکل ۶ آورده شده است. با توجه به این جدول، بیشینه انحراف متغیرها برای تمام نواحی تحت کنترل پیشنهادی از مقادیر متناظر تحت کنترل پیش‌بین مدل استاندارد و کنترل PID کمتر است. به عنوان مثال، بیشینه انحراف فرکانس، تحت کنترل پیشنهادی نسبت به کنترل پیش‌بین مدل استاندارد و کنترل PID به ترتیب 31.36% و 19.92% برای ناحیه اول، 23.99% و 17.93% برای ناحیه دوم و 29.64% و 7.94% برای ناحیه سوم کاهش یافته است. سرعت پاسخ متغیرهای انحراف فرکانس و خطای کنترل ناحیه، تحت کنترل پیشنهادی برای تمام نواحی نیز بهتر از مقادیر متناظر تحت کنترل‌های مورد مقایسه است. به عنوان مثال، زمان نشست متغیر خطای کنترل ناحیه تحت کنترل پیشنهادی نسبت به کنترل پیش‌بین مدل استاندارد و کنترل PID به ترتیب 44.84% و 61.39% برای ناحیه اول، 9.16% و 26.99% برای

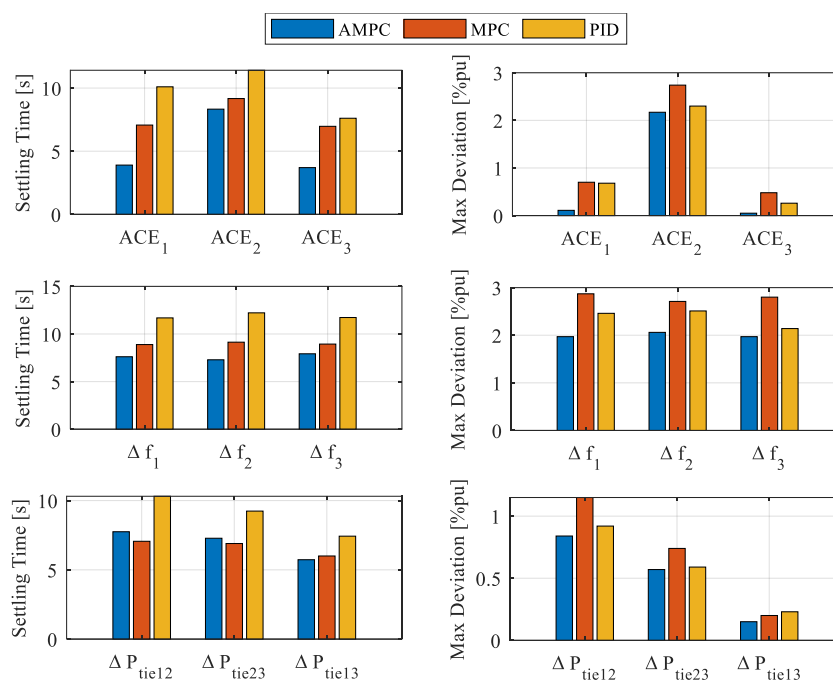
ناحیه دوم و 46.92% و 51.38% برای ناحیه سوم بهبود یافته است. همچنین، با وجود اغتشاش بار در ناحیه دوم، متغیر خروجی در این ناحیه نسبت به ناحیه‌های اول و سوم برای تمام کنترل‌کننده‌های مورد بررسی دارای بیشینه انحراف بزرگتری است. از آنجایی که می‌توان قید نرخ تولید را به‌طور خودکار در محاسبه سیگنال کنترل پیش‌بین مدل جابدهی کرد، کنترل پیشنهادی و کنترل پیش‌بین مدل استاندارد دارای سرعت پاسخ بهتر در مقایسه با کنترل PID هستند؛ به این ترتیب، عملکرد کنترل PID بیشتر از عملکرد کنترل پیشنهادی و کنترل پیش‌بین مدل استاندارد از قید نرخ تولید متاثر می‌شود.

کنترل پیش‌بین مدل تطبیقی نسبت به کنترل پیش‌بین مدل استاندارد از نظر کاهش بار محاسباتی نیز بهتر عمل می‌کند. زمان متوسط اجرای شبیه‌سازی برای کنترل PID برابر با 0.6511 s، برای کنترل پیش‌بین مدل استاندارد برابر با 1.7308 s و برای کنترل پیش‌بین مدل تطبیقی پیشنهادی برابر با 0.7216 s به طول می‌انجامد. به این ترتیب می‌توان گفت که بار محاسباتی کنترل پیش‌بین مدل تطبیقی پیشنهادی، با توجه به استفاده از مدل داخلی ساده شده و توابع لاگر در تخمین دنباله کنترلی، به اندازه قابل توجهی نسبت به کنترل پیش‌بین مدل استاندارد کاهش می‌یابد و دارای مقداری نزدیک به کنترل متداول PID است.



شکل ۵: انحراف فرکانس، خطای کنترل ناحیه و انحراف توان خط ارتباطی در سناریوی ۱ (پارامترهای ثابت نامی).
 جدول ۲: مقایسه عددی کنترل‌های مورد بررسی در سناریوی ۱ (پارامترهای ثابت نامی).

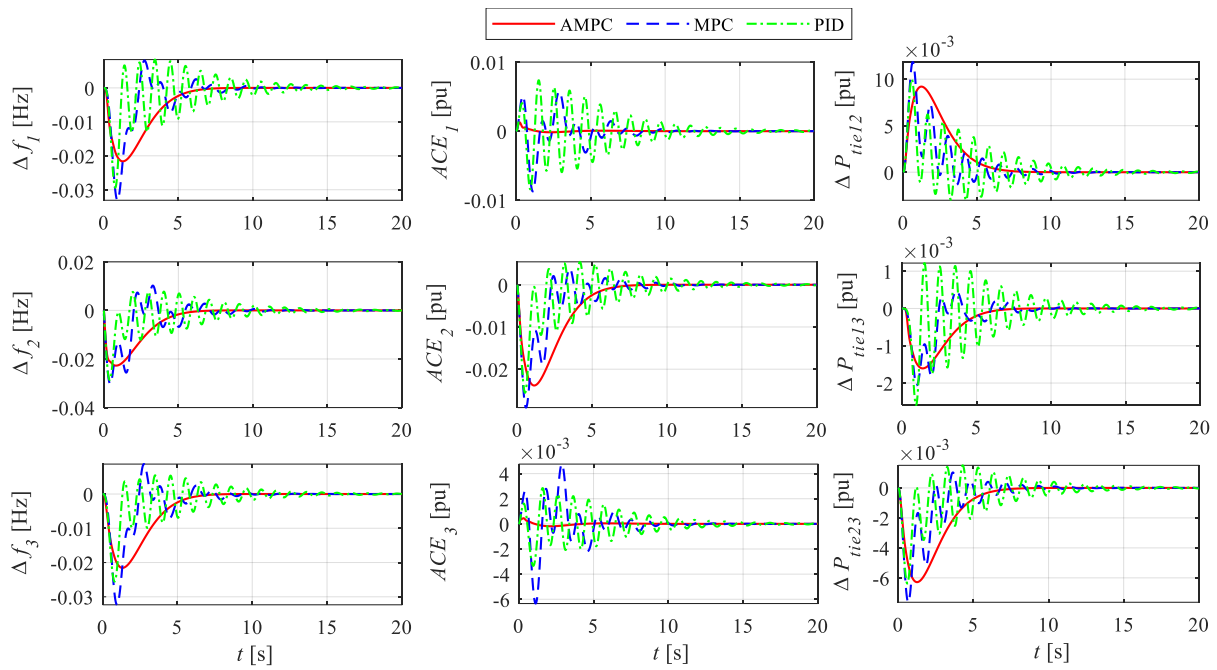
Variable	Settling Time [s]			Max. Deviation $\times 10^2$ [pu]		
	AMPC	MPC	PID	AMPC	MPC	PID
Δf_1	7.61	8.89	11.69	1.97	2.87	2.46
ACE_1	3.90	7.07	10.10	0.11	0.70	0.68
$\Delta P_{tie_{12}}$	7.75	7.06	10.31	0.84	1.15	0.92
Δf_2	7.29	9.14	12.22	2.06	2.71	2.51
ACE_2	8.33	9.17	11.41	2.17	2.74	2.30
$\Delta P_{tie_{23}}$	7.28	6.90	9.24	0.57	0.74	0.59
Δf_3	7.92	8.94	11.73	1.97	2.80	2.14
ACE_3	3.70	6.97	7.61	0.05	0.48	0.26
$\Delta P_{tie_{13}}$	5.73	6.00	7.43	0.15	0.20	0.23



شکل ۶: نمودار میله‌ای متغیرهای خطای کنترل ناحیه، انحراف فرکانس و انحراف توان خط ارتباطی در سناریوی ۱ (پارامترهای ثابت نامی). زمان نشست (چپ)، بیشینه انحراف (راست).

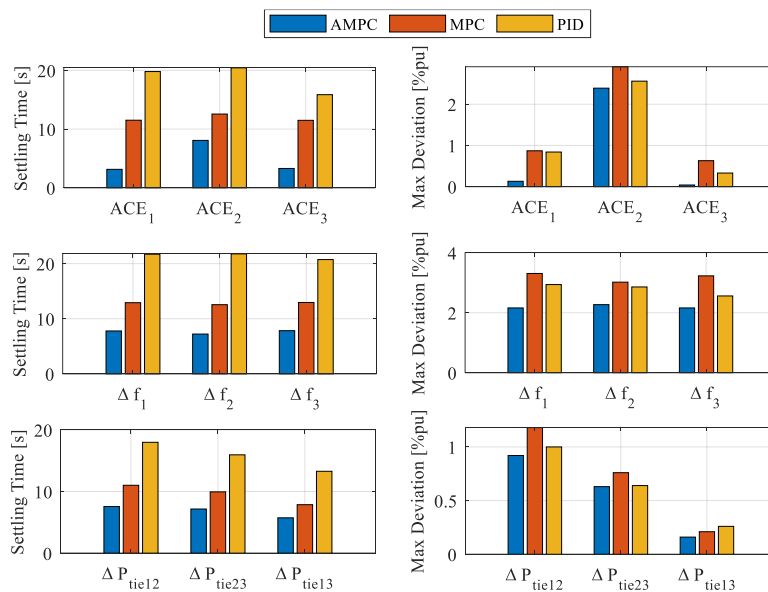
۴-۲- سناریوی ۲: بررسی عملکرد با تغییرات پارامتری

در این بخش، عملکرد مقاوم کنترل پیش‌بین تطبیقی پیشنهادی در دفع اغتشاش بار پله‌ای در ناحیه ۲ در حضور تغییرات پارامترهای سیستم مورد بررسی قرار گرفته است. با توجه به ابعاد و پیچیدگی سیستم‌های قدرت فعلی، وجود عواملی از قبیل اغتشاشات پیوسته بار و خطاهای سیستم سبب تغییر در پارامترهای مدل می‌شود. تضمین کیفیت توان و پایداری سیستم به مشخصه مقاومت کنترل‌کننده بار-فرکانس بستگی دارد. برای نشان دادن مقاومت کنترل پیشنهادی، فرض شده است که در زمان صفر به میزان 20% تغییرات در پارامترهای سیستم بوجود آمده است، به‌طوری‌که پارامترهای ثابت زمانی ژنراتور و ثابت زمانی توربین به 20% بیشتر از مقادیر اولیه و پارامترهای لختی سنکرون و ضریب میرایی به 20% کمتر از مقادیر اولیه تغییر کرده است. افزایش ثابت‌های زمانی و کاهش میرایی سیستم سبب کاهش پایداری ذاتی سیستم قدرت و در نتیجه چالش برای سیستم کنترل می‌شود. نتایج شبیه‌سازی مقایسه‌ای در شکل ۷، جدول ۳ و شکل ۸ آورده شده است. کنترل پیش‌بین مدل تطبیقی پیشنهادی با توجه به شناسایی آنلاین پارامترهای مدل قادر به حفظ عملکرد دفع اغتشاش بار با وجود تغییرات پارامتری است ولی عملکرد کنترل پیش‌بین مدل استاندارد و کنترل PID با افت قابل‌ملاحظه‌ای مواجه می‌شود. سرعت پاسخ حالت ماندگار سیستم را می‌توان به مشخصه مقاومت و قابلیت کنترل‌کننده در دفع اغتشاش بار ارتباط داد. به عنوان مثال، زمان نشست متغیر خطای کنترل ناحیه تحت کنترل پیشنهادی نسبت به کنترل پیش‌بین مدل استاندارد و کنترل PID به ترتیب 72.72% و 84.14% برای ناحیه اول، 35.75% و 60.50% برای ناحیه دوم و 71.37% و 79.26% برای ناحیه سوم بهبود یافته است. با توجه به این اعداد و نوسانات زیاد با میرایی ضعیف در پاسخ کنترل PID می‌توان دریافت که میزان افت عملکرد کنترل PID نسبت به کنترل پیش‌بین استاندارد در حضور تغییرات پارامتری بیشتر است؛ کنترل پیش‌بین مدل استاندارد در ابتدای هر گام کنترلی، اطلاعات به‌روزرسانی شده متغیرهای حالت فعلی سیستم را در قابل شرایط اولیه دریافت می‌کند و در نتیجه می‌تواند تا حدی اثر تغییرات پارامتری را پوشش دهد.



شکل ۷: انحراف فرکانس، خطای کنترل ناحیه و انحراف توان خط ارتباطی در سناریوی ۲ (20% تغییرات پارامتری).
 جدول ۳: مقایسه عددی کنترل‌های مورد بررسی در سناریوی ۲ (20% تغییرات پارامتری).

Variable	Settling Time [s]			Max. Deviation $\times 10^2$ [pu]		
	AMPC	MPC	PID	AMPC	MPC	PID
Δf_1	7.78	12.91	21.71	2.16	3.31	2.94
ACE_1	3.14	11.51	19.80	0.13	0.87	0.84
$\Delta P_{tie_{12}}$	7.56	11.01	17.98	0.92	1.18	1.00
Δf_2	7.22	12.55	21.77	2.27	3.02	2.86
ACE_2	8.07	12.56	20.43	2.39	2.91	2.56
$\Delta P_{tie_{23}}$	7.16	9.94	15.95	0.63	0.76	0.64
Δf_3	7.83	12.95	20.74	2.16	3.23	2.56
ACE_3	3.29	11.49	15.86	0.04	0.63	0.33
$\Delta P_{tie_{13}}$	5.73	7.86	13.28	0.16	0.21	0.26



شکل ۸: نمودار میله‌ای متغیرهای خطای کنترل ناحیه، انحراف فرکانس و انحراف توان خط ارتباطی در سناریوی ۲ (20% تغییرات پارامتری). زمان نشست (چپ)، بیشینه انحراف (راست).

۵- نتیجه‌گیری

در این مقاله یک کنترل پیش‌بین مدل تطبیقی برای مساله کنترل بار-فرکانس یک سیستم قدرت بهم‌پیوسته چند ناحیه‌ای در حضور تغییرات پارامتری پیشنهاد شد. یک مدل ساده شده ARX از سیستم قدرت اصلی به کمک تخمینگر مدل چند جمله‌ای بازگشتی در ابتدا به‌طور آفلاین تخمین زده شد و سپس در قالب مدل پیش‌بینی کنترل‌کننده، در هر گام زمانی به‌طور آفلاین به‌روزرسانی می‌شود. یک فیلتر کالمن خطی تغییرپذیر با زمان وظیفه رویت متغیرهای سیستم را بر عهده دارد. عملکرد کنترل پیشنهادی با عملکرد کنترل پیش‌بین مدل استاندارد و کنترل PID در دفع اغتشاش بار پله‌ای در سیستم قدرت با پارامترهای ثابت نامی و پارامترهای متغیر با زمان مورد مقایسه قرار گرفت. نتایج حاصل از خروجی شبیه‌سازی را می‌توان به این ترتیب جمع‌بندی کرد: (۱) با توجه به شناسایی یک مدل کاهش مرتبه یافته و استفاده از توابع لاگر برای تخمین دنباله کنترلی آتی، بار محاسباتی کنترل پیش‌بین مدل تطبیقی پیشنهادی به اندازه زیادی در مقایسه با کنترل پیش‌بین مدل استاندارد کاهش می‌یابد و قابل مقایسه با کنترل متداول PID است. (۲) با توجه به تخمین آفلاین پارامترهای مدل و به‌روزرسانی مدل پیش‌بینی در هر گام کنترلی، عملکرد پاسخ دینامیکی (از نظر زمان نشست و بیشینه انحراف) کنترل پیشنهادی در مقایسه با کنترل پیش‌بین مدل استاندارد و به ویژه کنترل PID به‌طور قابل ملاحظه‌ای بهبود می‌یابد. بنابراین، کنترل پیش‌بین مدل تطبیقی پیشنهادی می‌تواند جایگزین مناسبی برای کنترل متداول PID برای مساله بار-فرکانس سیستم‌های قدرت پیچیده فعلی باشد. از آنجایی که بکارگیری منابع انرژی تجدیدپذیر در سیستم‌های قدرت فعلی رو به رشد است، کارهای آتی به توسعه یک مدل جامع‌تر از سیستم قدرت با در نظر گرفتن منابع انرژی تجدیدپذیر متداول از قبیل مزارع بادی و خورشیدی و بررسی اثر دینامیک‌ها و خروجی نوسانی این واحدها روی عملکرد سیستم کنترل بار-فرکانس اختصاص دارد.

مراجع

- [1] D. Jiao, C. Shao, B. Hu, K. Xie, C. Lin, Z. Ju, "Age-of-information-aware PI controller for load frequency control," *Prot. Control Mod. Power Syst.*, vol. 8, no. 3, p. 631–645, 2023.
- [2] F. Milano, F. Dörfler, G. Hug, D.J. Hill, G. Verbič, "Foundations and challenges of low-inertia systems," in *Power Syst. Computation Conf*, Dublin, Ireland, 2018.
- [3] Y. Zhao, T. Liu, D. J. Hill, "Data-driven adaptive predictive frequency control for power systems with unknown and time-varying inertia," *Electric Power Systems Research*, vol. 234, p. 110815, 2024.
- [4] S. Tripathi, V.P. Singh, N. Kishor, A.S. Pandey, "Load frequency control of power system considering electric Vehicles' aggregator with communication delay," *Int J Electr Power Energy Syst*, vol. 145, p. 108697, 2023.
- [5] L. Jin, et al., "Delay-dependent stability of load frequency control with adjustable computation accuracy and complexity," *Control Eng Pract.*, vol. 135, p. 105518, 2023.
- [6] B. Khokhar, K. Pal Singh Parmar, "Impact analysis of capacitive energy storage integration on load frequency control performance of microgrids employing a new dual-stage controller," *ISA Transactions*, 2025.
- [7] J. Ansari, A. Abbasi, B. Bahmani Firouzi, "Decentralized LMI-based event-triggered integral sliding mode LFC of power systems with disturbance observer," *Int J Electr Power Energy Syst.*, vol. 138, p. 107971, 2022.
- [8] Z. Ming, H. Zhang, J. Sun, "Self-triggered load frequency control using T-S fuzzy ADP method for unknown power systems," *ISA Transactions*, 2025.

- [9] P. Samal, N. Nayak, A. Satapathy, S. Kumar Bhuyan, "Load frequency control in renewable based micro grid with Deep Neural Network based controller," *Results in Engineering*, vol. 25, p. 103554, 2025.
- [10] J. Li, T. Zhou, "Fully autonomous load frequency control for integrated energy system with massive energy prosumers using multi-agent deep meta reinforcement learning," *Renewable and Sustainable Energy Reviews*, vol. 213, p. 115489, 2025.
- [11] W. Qian, D. Lu, Y. Wu, "Adaptive memory event-triggered load frequency control for multiarea power systems with non-ideal communication channel," *Electric Power Systems Research*, vol. 245, p. 111596, 2025.
- [12] X. H Linyun, L. Hao, W. Jie, "LMI based robust load frequency control for time delayed power system via delay margin estimation," *Int. J. Electr. Power Ener. Syst.*, vol. 100, pp. 91-103, 2018.
- [13] G. I.Y. Mustafa, MD. Masum, MD. Siam, "A new model-free control for load frequency control of interconnected power systems based on nonlinear disturbance observer," *Energy Reports*, vol. 25, p. 4998–5008, 2024.
- [14] V. Vasuki Gautam, R. Loka, A. M. Parimi, "Cubature Kalman filter and linear quadratic regulator for load frequency control," *Electric Power Systems Research*, vol. 222, p. 109509, 2023.
- [15] S. Mokhtari, K. K. Yen, "Dynamic state estimation with additive noise for load frequency control using bilateral fuzzy adaptive unscented Kalman filter," *Electric Power Systems Research*, vol. 220, p. 109363, 2023.
- [16] A.M. Ersdal, L. Imsland & K. Uhlen, "Model predictive load-frequency control," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 777-785, 2016.
- [17] N. Zhao, et al., "Model predictive based frequency control of power system incorporating air-conditioning loads with communication delay," *Int J Electr Power Energy Syst.*, vol. 138, p. 107856, 2022.
- [18] Y Zheng, J Zhou, Y Xu, Y Zhang & Z Qian, "A distributed model predictive control based load frequency control scheme for multi-area interconnected power system using discrete-time Laguerre functions," *ISA Transactions*, vol. 68, pp. 127-140, 2017.
- [19] W. Liu, Y. Liu, "Hierarchical model predictive control of wind farm with energy storage system for frequency regulation during black-start," *Int J Electr Power Energy Syst.*, vol. 119, p. 105893, 2020.
- [20] A. Oshnoei, M. Kheradmandi, R. Khezri, A. Mahmoudi, "Robust model predictive control of gate-controlled series capacitor for LFC of power systems," *IEEE Trans Ind Inf.*, vol. 17, no. 7, p. 4766–4776, 2021.
- [21] G. Cai, et al., "Data-driven predictive based load frequency robust control of power system with renewables," *Electrical Power and Energy Systems*, vol. 154, p. 109429, 2023.
- [22] Z. An, et al., "Tube-based MPC strategy for load frequency control of multi-area interconnected power system with HESS," *Journal of Energy Storage*, vol. 99, p. 113340, 2024.
- [23] C. Huang, et al., "DMPC-based load frequency control of multi-area power systems with heterogeneous energy storage system considering SoC consensus," *Electric Power Systems Research*, vol. 228, p. 110064, 2024.
- [24] P. Wang, et al., "A MPC-based load frequency control considering wind power intelligent forecasting," *Renewable Energy*, vol. 244, p. 122636, 2025.
- [25] M. Ma, J. Cui, X. Liu, K. Y. Lee, "Distributed Economic Model Predictive Load Frequency Control for the Multiarea Interconnected Power System With WTs," *IEEE Systems Journal*, vol. 18, no. 3, pp. 1629-1638, 2024.
- [26] P.A. Gbadega, A.K. Saha, "Load Frequency Control of a Two-Area Power System With a Stand-Alone Microgrid Based on Adaptive Model Predictive Control," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 6, pp. 7253-7263, 2021.

- [27] W. Wang, et al., "A novel adaptive model predictive frequency control using unscented Kalman filter," *Electric Power Systems Research*, vol. 213, p. 108721, 2022.
- [28] W. Wang, et al., "Adaptive Model Predictive-Based Load Frequency Controller using Unscented Kalman Filter," in *IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia)*, Brisbane, Australia, 2021.
- [29] S. Jiang, et al., "Feedback-Based Setpoint Mechanism in Model Predictive Load Frequency Control," *IEEE Transactions on Power Systems*, vol. 39, no. 4, pp. 6115 - 6118, 2024.
- [30] T.H. Mohamed, J. Morel, H. Bevrani & T. Hiyama, "Model predictive based load frequency control_design concerning wind turbines," *International Journal of Electrical Power & Energy Systems*, vol. 43, no. 1, pp. 859-867, 2012.
- [31] M. Bujarbaruah, X. Zhang, M. Tanaskovic, F. Borrelli, "Adaptive stochastic MPC under time-varying uncertainty," *IEEE Trans. Automat. Control*, vol. 66, no. 6, p. 2840–2845, 2021.
- [32] J.B. Jorgensen, J.K. Huusom, J.B. Rawlings, "Finite horizon MPC for systems in innovation form," in *50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, FL, USA, 2011.
- [33] N. Assimakis, M. Adam, "Kalman filter Riccati equation for the prediction, estimation, and smoothing error covariance matrices," *ISRN Comput Math*, vol. 2013, no. 1, pp. 1-7, 2013.
- [34] L. Ljung, *System Identification—Theory for the User*, Upper Saddle River, New Jersey: Prentice Hall PTR, 1999.
- [35] D. Ioan, G.Z. Landau, *Adaptive Control*, London: Springer-Verlag, 1997.
- [36] J.L. Garriga, M. Soroush, "Model predictive control tuning methods: a review," *Industrial & Engineering Chemistry Research*, vol. 49, no. 8, pp. 3505-3515, 2010.
- [37] A.L. Bruce, A. Goel, D.S. Bernstein, "Convergence and consistency of recursive least squares with variable-rate forgetting," *Automatica*, vol. 119, pp. 1-6, 2020.
- [38] Adel A.B. Rehiara, H. Chongkai, Y. Sasaki, N. Yorino & Y. Zoka, "An Adaptive Internal Model for Load Frequency Control Using Extreme Learning Machine," *TELKOMNIKA*, vol. 16, no. 6, pp. 1-6, 2018.

نشریه سامانه‌های غیرخطی در مهندسی برق نشریه‌ای علمی-پژوهشی است که جدیدترین نتایج پژوهش‌های بنیادی، کاربردی و توسعه‌ای در زمینه سامانه‌های غیرخطی مهندسی برق را چاپ می‌کند. از میان مباحث موردنظر این نشریه می‌توان به موارد زیر اشاره کرد:

- نور غیرخطی و کاربردهای آن
- سامانه‌های دینامیکی غیرخطی و کاربردهای آنها
- ماشین‌های الکتریکی و سامانه‌های قدرت شامل شبکه‌های تولید، انتقال و توزیع
- مدل‌سازی، شناسایی، تخمین، کنترل، بهینه‌سازی و پیاده‌سازی عملی الگوریتم‌های تحمل‌پذیر عیب
- آشوب و سنکرون‌سازی
- سامانه‌های هابیرید، سویچینگ و چندعامل
- رباتیک و بینایی ماشین
- شبکه‌های مخابراتی، سیستم‌های الکترونیکی و مبدل‌های داده
- پردازش سیگنال‌ها، تصاویر و ویدئو و اسناد، بازشناسی الگو

فرمت نگارش مقالات

مقالات در حال حاضر فقط به زبان فارسی پذیرفته می‌شوند. مقالات ارسالی باید به ترتیب شامل بخش‌های زیر باشند:
عنوان (فارسی)، نام نویسندگان همراه با نام مؤسسه متبوع هر نویسنده، چکیده (فارسی)، واژه‌های کلیدی (فارسی)، مقدمه، بدنه مقاله، نتیجه‌گیری، سپاسگزاری (در صورت وجود)، پیوست (در صورت وجود)، مراجع، شکل‌ها، جداول و تصاویر.
شماره‌گذاری بخش‌ها می‌بایست مطابق شماره‌گذاری مرسوم عددی باشد.

شیوه تدوین (مقاله فقط فارسی)

متن مقالات شامل چکیده، بدنه مقاله، مراجع و زیرنویسها باید با فونت ۱۲ B zar و با فاصله double میان خطوط، در صفحات A4 یک‌رو و تک ستونی با استفاده از نرم‌افزار Word 2010 تهیه شود. علاوه بر این، فایل PDF مقاله (بدون ذکر نام نویسندگان و مؤسسه متبوع و آدرس ایمیل) نیز می‌بایست ارسال شود.

آدرس نویسندگان

آدرس پستی کامل همه نویسندگان همراه با شماره تلفن و دورنگار (فکس) و نشانی الکترونیکی (E-Mail) نویسنده عهده‌دار مکاتبات (که الزاماً می‌بایست عضو هیأت علمی باشند) در برگه مستقلی چاپ و به همراه مقاله ارسال گردد.

چکیده

هر مقاله باید شامل عنوان (فارسی)، چکیده مقاله (فارسی) در حداکثر ۲۰۰ واژه، واژه‌های کلیدی (فارسی) در حداکثر ۵ واژه باشد.

شکل‌ها و تصاویر

در هنگام ارسال مقاله به منظور داوری نیازی به ارسال اصل شکل‌ها و تصاویر نیست، ولی رونوشت ارسالی باید واضح باشد. پس از تأیید مقاله، ارسال اصل شکل‌ها و تصاویر (با فرمت Tiff) برای چاپ مقاله ضروری است. شکل‌ها و تصاویر نباید قابل ویرایش باشند.

مراجع

به کلیه مراجع (به تعداد حداقل ۱۰ مرجع) باید در متن ارجاع داده شده باشد. مراجع باید تابع استاندارد IEEE بوده و با شماره، مشخص گردند و جزئیات آنها به شرح زیر در پایان مقاله ظاهر گردد:

مقالات: شماره مرجع، نام خانوادگی و نام کامل همه نویسندگان، "عنوان مقاله"، نام کامل مجله یا کنفرانس، سال انتشار یا تاریخ برگزاری، شماره مجله یا شماره جلد، شماره صفحات.

کتابها: شماره مرجع، نام خانوادگی و نام کامل همه نویسندگان، عنوان کتاب، نام مترجم، نام کامل ناشر، سال انتشار.

واحدها

کلیه مقالات باید از واحد استاندارد SI در تمام بخش‌های مقاله استفاده نمایند. در کنار واحد SI می‌توان از واحد انگلیسی در داخل پیرانتز نیز استفاده نمود.

تعداد صفحات مقاله

حداکثر تعداد صفحات مقالات در هنگام چاپ، ۱۲ صفحه است. برای چاپ صفحات بیشتر و یا رنگی، لازم است هزینه‌ای برای هر صفحه به حساب نشریه واریز شود.

فرآیند ارسال مقاله

مقالات قابل چاپ در نشریه شامل مقالات کامل پژوهشی و مقالات کوتاه پژوهشی هستند. مقالات ارسالی نباید در هیچ نشریه داخلی و یا خارجی چاپ شده باشند و یا در حال داوری باشند. لازم است نسخه اولیه مقاله به صورت تک ستونی در دو نسخه (یک نسخه با نام نویسندگان و مؤسسه متبوع نویسندگان با فرمت docx و یک نسخه بدون نام نویسندگان و مؤسسه متبوع نویسندگان و آدرس ایمیل با فرمت pdf. به منظور ارسال به داوران) از طریق سایت اینترنتی نشریه به آدرس <http://jnsee.sut.ac.ir> برای داوری ارسال شود. داوری به شکل Double Blind Review است و نسخه اولیه مقالات به داوران متخصص ارسال می‌شود. در پایان، تأیید یا رد هر مقاله توسط هیأت تحریریه نشریه انجام خواهد شد. سردبیر نشریه نتیجه داوری را برای نویسنده عهده‌دار مکاتبات ارسال خواهد کرد. در صورتی که نیاز به تصحیح مقاله باشد، تصحیحات باید منحصراً محدود به موارد اصلاحی ذکر شده باشند. در سایر موارد، نویسنده لازم است سردبیر را در جریان هر گونه تغییر و یا تصحیح دیگری قرار دهد. در هر صورت، مسئولیت صحت و سقم مطالب بر عهده نویسنده/نویسندگان مقاله خواهد بود.

در صورتی که مقاله برای چاپ پذیرفته شود، لازم است اصل مقاله مطابق فرمت استاندارد نشریه از طریق سایت نشریه ارسال شود. پس از چاپ، یک نسخه از نشریه به هر یک از نویسندگان اهدا خواهد شد.

حق کپی‌رایت

در صورت پذیرش مقاله برای چاپ، نویسندگان لازم است فرم انتقال حق انتشار آن به "دانشگاه صنعتی سهند" را تکمیل و به همراه اصل مقاله ارسال نمایند. نویسندگان لازم است موافقت کتبی دارندگان حق کپی بخش‌هایی از مقاله که از مراجع و منابع دیگر نسخه‌برداری شده است را دریافت و به دفتر مجله ارسال نمایند.

بدینوسیله از کلیه اساتید و پژوهشگران برای ارسال مقالات خویش در این نشریه دعوت به عمل می‌آید. خواهشمند است مقالات خود را از طریق سایت نشریه به آدرس <http://jnsee.sut.ac.ir/> ارسال فرمایید.