

Monocular Visual Odometry

فرهاد دلیرانی

تعیین مکان دقیق یک وسیله نقلیه یک چالش اساسی در حوزه‌ی ربات‌های متحرک است. یک ربات برای اینکه به ناوبری اتوماتیک دست بیاورد باید اطلاعات دقیقی از مکانش در طول زمان داشته باشد. به همین دلیل انواع سیستم‌ها، سنسورها و روش‌ها برای تعیین مکان ربات‌ها و وسایل نقلیه متحرک ایجاد شده‌اند.

بعضی از این روش‌ها و سنسورها :

GPS □

Wheel Odometry □

Laser Sensors □

Sonar □

INS □

Visual Odometry □

- ▶ کلمه **Odometry** از دو کلمه یونانی **Hodos** و **Metron** گرفته شده است که آن دو کلمه در یونانی معنی **سفر و اندازه‌گیری** می‌دهند. ربات‌های متحرک با استفاده از اطلاعاتی که از سنسورهای مختلف خود می‌گیرند، **موقعیت مکانی** خود را در **طول زمان** به صورت نسبی از نقطه شروع محاسبه می‌کنند که به این فرآیند **Odometry** می‌گویند.
- ▶ در صورتی که در تعیین مکان ربات متحرک فقط از **جریانی از تصاویر** که از **یک یا چند دوربین** متصل به ربات گرفته می‌شوند استفاده شود به آن **Visual Odometry** می‌گویند.
- ▶ **تصاویر** حاوی اطلاعات **کافی** و با معنی (رنگ، بافت، شکل و ...) هستند که می‌توانند برای تعیین مکان دوربین در طول زمان استفاده شوند.

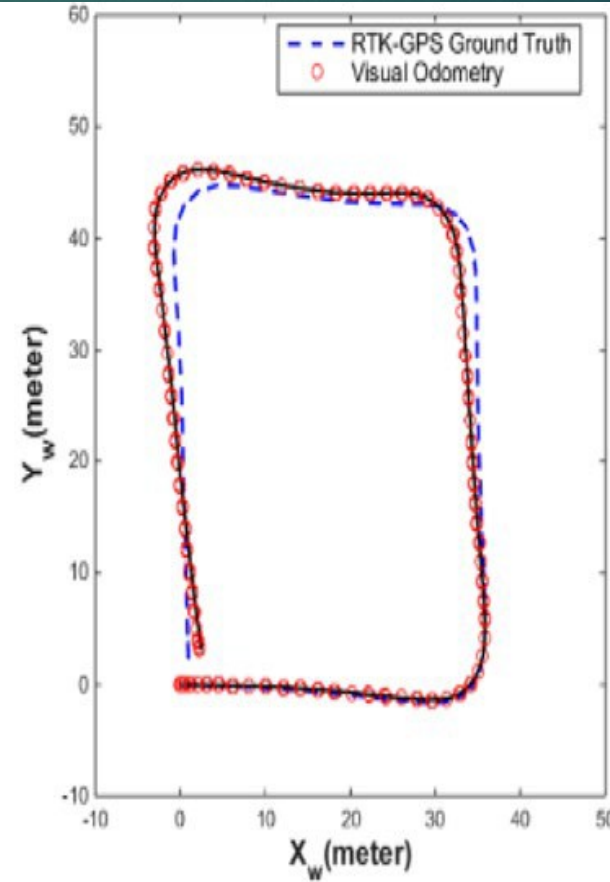


Visual Odometry



Input

Image sequence
(video stream)



Output

Camera trajectory
(translation and rotation)

مقاله‌های مطالعه شده برای کار مطالعاتی

1

2

3

4

5

6

Monocular Visual Odometry Based on Optical Flow and Feature Matching

Cheng Chuanqi^{1,2}, Hao Xiangyang^{1,2}, Zhang Zhenjie^{1,2}, Zhao Mandan^{1,2}

1. School of Navigation and Aerospace Engineering, Information Engineering University, Zhengzhou 450001, China

2. Beidou Navigation Technology Collaborative Innovation Center of Henan, Zhengzhou 450001, China

E-mail: legend3q@163.com xiangyanghao2004@163.com zzxiaodao@126.com mandanzhao@163.com

Abstract: To solve the problem of real-time precise localization in CPS-denied places, a monocular visual odometry method based on optical flow tracking and feature matching is proposed. To speed up traditional pose estimation algorithm, the image sequences are classified into key frames and non-key frames. The conventional pipeline of feature detection and matching is utilized to process key frames, while utilizing Lucas Kanade optical flow to track the correspondences in non-key frames. To improve the robustness of the visual odometry method, a RANSAC-based outlier rejection scheme is applied in the phase of pose estimation. Moreover, a Kalman Filter based on the dynamic equation is designed to optimize the pose estimation. Experimental results demonstrate that proposed method can acquire the high accuracy of feature matching, while highlighting the real-time performance of optical flow tracking, which can meet the needs of real-time accurate localization in cities.

Key Words: Monocular Visual Odometry, Optical Flow, RANSAC, Feature Matching, Kalman Filter

1 INTRODUCTION

Precise fully autonomous positioning is of great importance

2017 29th Chinese Control
And Decision Conference
(CCDC)

Publisher: IEEE

sensors to increase the accuracy of the results and reduce drift error (e.g., [6], [7], [8], [9], [10]). While Li et. al. [6] used an IMU, Agrawal et. al. [10] used GPS and wheel encoders, to seek the optimal performance. Moreover, methods making assumptions about the camera's motion have been developed [11]. Kitt et. al. proposed an algorithm for egomotion estimation in all six degrees of freedom using a fully calibrated stereo camera rig [12].

Compared to the methods combining other sensors, our approach mainly focuses on estimating the pose solely



بخش یک

ادومتری بصری: انواع، روش‌ها، چالش‌ها و کاربردها

بخش دو

Monocular Visual اساس و پایه روش
Odometry

بخش سه

Monocular Visual برخی از روش‌های
Odometry

ادومتری بصری: انواع، روش‌ها، چالش‌ها و کاربردها

Review of visual odometry: types, approaches, challenges, and applications



Mohammad O. A. Aqel^{1*}, Mohammad H. Marhaban², M. Iqbal Saripan³ and Napsiah Bt. Ismail⁴

*Correspondence:
aqel2001@hotmail.com
¹ Department of Engineering,
Faculty of Engineering
and Information Technology,
Al-Azhar University-Gaza,
Gaza, Palestine
Full list of author information
is available at the end of the
article

Abstract

Accurate localization of a vehicle is a fundamental challenge and one of the most important tasks of mobile robots. For autonomous navigation, motion tracking, and obstacle detection and avoidance, a robot must maintain knowledge of its position over time. Vision-based odometry is a robust technique utilized for this purpose. It allows a vehicle to localize itself robustly by using only a stream of images captured by a camera attached to the vehicle. This paper presents a review of state-of-the-art visual odometry (VO) and its types, approaches, applications, and challenges. VO is compared

Visual Odometry

Part I: The First 30 Years and Fundamentals

By Davide Scaramuzza and Friedrich Fraundorfer

سنسورها و روش‌های مورد استفاده برای تعیین مکان

► برای پی‌بردن به اهمیت و دلیل استفاده از Visual Odometry، در ابتدا به صورت خلاصه به روش‌ها و سنسورهای دیگری می‌پردازیم.



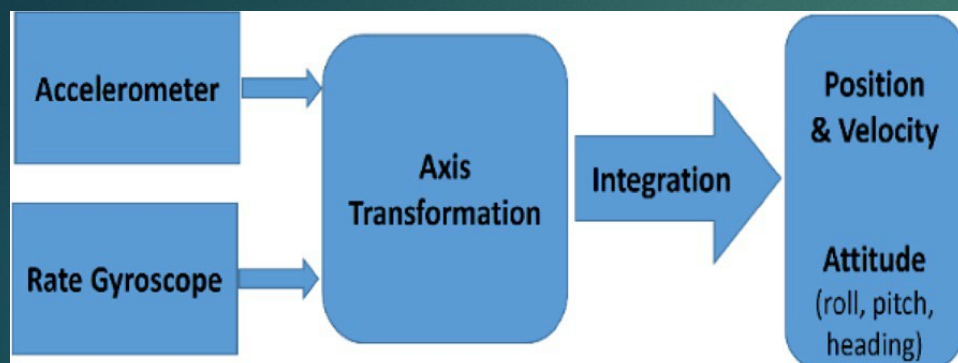
► **Wheel Odometry**

► این روش در ربات‌های متحرکی که با استفاده از چرخ حرکت می‌کنند استفاده می‌شود. این روش برای تعیین مکان نسبی استفاده می‌شود. سرخوردن و تغییر شکل چرخ باعث ایجاد خطا در مسیر می‌شود. خطا به صورت تجمعی است.

► این روش ساده و ارزان است و برای استفاده در مقطع‌های زمانی کوتاه مناسب است.

سنسورها و روش‌های مورد استفاده برای تعیین مکان

INS (inertial Navigation System) ➤



a



b

➤ INS یک روش **مکانیابی نسبی** است که مکان و جهت یک ربات متحرک نسبت به نقطه آغاز را مشخص می‌کند.

➤ در این روش از یک کامپیوتر، **سنسور شتاب‌سنج** و **سنسور ژيروسکوپ** استفاده می‌شود و به طور پیوسته جهت، مکان و شتاب وسیله سنجیده و محاسبه می‌شود.

➤ در این وسیله محدودیتی در **نوع ربات** متحرک نداریم، ربات می‌تواند یک ربات هوایی، زمینی و یا آبی باشد.

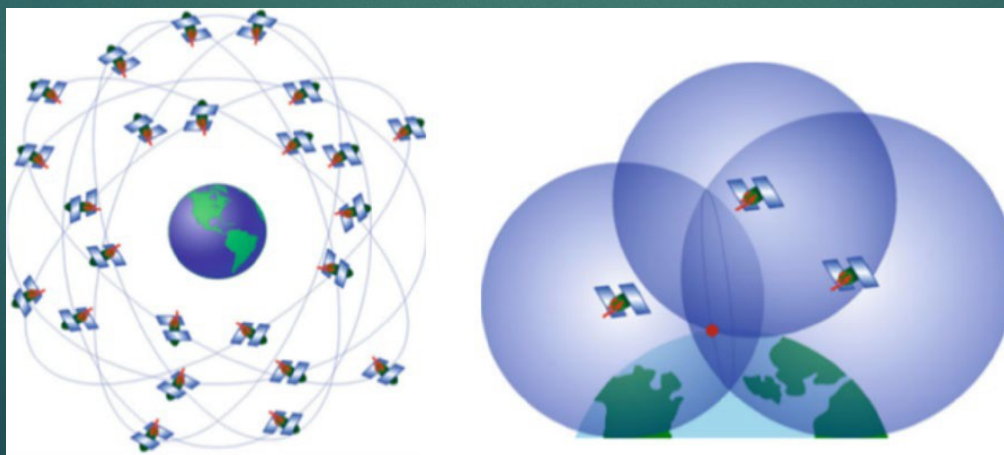
➤ به دلیل **تجمعی بودن خطا** برای استفاده در **بازه‌های زمانی طولانی** مناسب نیست.

➤ سنسورهای INS با دقت‌ها مختلفی موجود است. با زیاد شدن دقت، **قیمت** اینگونه سنسورها بسیار زیاد می‌شود.

سنسورها و روش‌های مورد استفاده برای تعیین مکان

GNSS(Global Navigation Satellite System) ➤

- یک نام کلی برای تمام سیستم‌های ناوبری جهانی بر اساس رادیوولست که شامل جی.پی.اس هم می‌شود.
- جی.پی.اس یک روش ناوبری بر اساس ماهواره‌ها است که اجازه تعیین موقعیت به هر شی که در مکانی سرپوشیده نباشد را می‌دهد.
- موقعیت مکانی جسم را می‌دهد و مانند روش‌های قبل موقعیت به صورت نسبی نیست. خطا تجمعی نیست.
- دقتش در حد چند متر است که برای بسیاری از استفاده‌ها مناسب نیست.



سنسورها و روش‌های مورد استفاده برای تعیین مکان

➤ سنسورهای سونار و لیزری

➤ این سنسورها از لیزر و موج‌های فراصوت برای تشخیص اشیا و محاسبه فاصله استفاده می‌کنند.

Sensor/technology	Advantages	Disadvantages
Wheel odometry	Simple to determine position/orientation Short term accuracy, and allows high sampling rates Low cost solution	Position drift due to wheel slippage Error accumulation over time Velocity estimation requires numerical differentiation that produces additional noise
INS	Provides both position and orientation using 3-axis accelerometer and gyroscope Not subject to interference outages	Position drift (position estimation requires second-order integral) Have long-term drift errors
GPS/GNSS	Provides absolute position with known value of error No error accumulation over time	Unavailable in indoor, underwater, and closed areas Affected by RF interference
Ultrasonic sensor	Provides a scalar distance measurement from sensor to object Inexpensive solution	Reflection of signal wave is dependent on material or orientation of obstacle surface Suffer from interference if multiple sensors are used Low angular resolution and scan rate
Laser sensor	Similar to sonar sensors but has higher accuracy and scan rate Return the distance to a single point (rangefinder) or an array of distances (scanner)	Reflection of signal wave is dependent on material or orientation of obstacle surface Expensive solution

سنسورها و روش‌های مورد استفاده برای تعیین مکان

دوربین

- ▶ اخیراً پژوهشگران بسیاری به سیستم‌های تعیین مکان براساس دوربین علامند شده‌اند. دلیل این امر رباست بودن و قابل اطمینان بودن این سیستم‌ها نسبت به بسیاری از روش‌ها و سنسورهای تعیین موقعیت است.
- ▶ تصویرهای گرفته شده با دوربین می‌توانند در محیط‌های سرباز و سرپوشیده برای تعیین موقعیت مورد استفاده قرار بگیرد.
- ▶ نسبت به بسیاری از سنسورها و روش‌ها، دوربین‌ها وسایلی بسیاری ارزان هستند که حجم بسیار زیادی از اطلاعات را در اختیار قرار می‌دهند.
- ▶ Visual Odometry یک روش افزایشی و برخط برای تعیین موقعیت‌های متحرک است که در لین کار از دنباله‌ای از تصاویر استفاده می‌کند. اودومتري بصری نسبت به بسیاری از روش‌ها زایلتر است و دقت نسبی بیشتری نسبت به آن‌ها دارد.

سنسورها و روش‌های مورد استفاده برای تعیین مکان

► **Visual Odometry** یک تعادل مناسب بین قیمت دقت قابل‌تکا بودن و پیچیدگی پیاده‌سازی دارد. با این حال در پردازش تصاویر معمولاً به محاسبات زیادی نیاز است محاسبات معمولاً شامل موردهای زیر است

۱. ثبت تصاویر
 ۲. استخراج یک یا چند ویژگی از تصاویرها (لبه، گوشه، خط و ...)
 ۳. پیدا کردن تطابق بین فریم‌های مختلف
 ۴. محاسبه موقعیت با بررسی تغییر مکان پیکسل‌ها در دو فریم
- الگوریتم‌های بینایی ماشین نسبت به شرایط محیطی که تصاویر در آن‌ها گرفته می‌شوند حساس‌اند. از این شرایط می‌توان به نورپردازی، بافت، تغییر شرایط نوری در طول زمان، حضور سایه، وجود آب و برف در صحنه اشاره کرد.

Sensor/technology	Advantages	Disadvantages
Optical camera	Images store a huge meaningful information Provide high localization accuracy Inexpensive solution	Requires image-processing and data-extraction techniques High computational-cost to process images

Visual Odometry قدمت



► Visual Odometry به فرآیند تعیین موقعیت یک عامل انسان وسیله نقلیه، ربات و ...) با استفاده از دنباله‌ای از تصاویر لاستکه توسط یک یا چند دوربین متصل به عامل گرفته می‌شوند روش ادومتری بصری یک روش افزایشی و برخط است.



ایده تعیین موقعیت اشیا با استفاده از تصاویر در ابتدا در سال ۱۹۸۰ توسط **Moravec** (موراوتس) مطرح شد.

در سال‌های ۱۹۸۰ تا ۲۰۰۰ عمدی پژوهش‌ها در حوزه ادومتری بصری توسط **ناسا** انجام شد که هدف آن پژوهش‌ها ارسال کاوشگر به **مریخ** در سال ۲۰۰۴ بود.

چرا از Visual Odometry استفاده می شود؟

- ▶ ادومتری بصری یک جایگزین ارزان تر و دقیق تر برای روش های دیگر ادومتری مانند ادومتری چرخ، جی.پی.اس، آی.ان.اس و ... است.
- ▶ خطای ادومتری بصری حدود ۰.۱ تا ۲ درصد است.
- ▶ این روش به خاطر تعادل بین قیمت، قابل اطمینان بودن و پیچیدگی پیاده سازی شناخته می شود.
- ▶ تصاویر حجم زیادی از اطلاعات را شامل می شوند که با استفاده از آن اطلاعات می توان موقعیت ربات را پیدا کرد.
- ▶ ادومتری بصری برخلاف بسیاری از روش ها بر اثر لغزش چرخ دچار خطا نمی شود.
- ▶ ادومتری بصری، وابسته به نوع خاصی از ربات ها نیست. ربات های زمینی، هوایی و آبی می توانند از آن استفاده کنند.
- ▶ همچنین این روش در محیط های سرپوشیده بر خلاف سیستم هایی مانند جی.پی.اس به خوبی کار می کند.
- ▶ ادومتری بصری را می توان با سایر روش ها به راحتی ترکیب کرد و به بیشترین دقت رسید.
- ▶ از تصاویری که برای ادومتری بصری استفاده می شود می توان برای سایر کارهای بینایی ماشین مانند تشخیص مانع، تشخیص انسان، تعامل با محیط و ... استفاده کرد.

چالش‌های Visual Odometry

- ▶ ادومتری بصری در محیط‌های سرپوشیده و بسته به خوبی مورد پژوهش قرار گرفته است و نتایج بسیار خوبی به دست آمده است ولی ادومتری بصری در محیط‌ها خارج از ساختمان، هنوز مسئله‌ای چالش برانگیز باقی مانده است.
- ▶ عوامل زیادی مانند زمین ناهموار، تلبش مستقیم خورشید، سایه، تغییرات در محیط مانند تغییرهای ناشی از وزش باد و ... تعیین مکان در محیط‌های خارج ساختمان را سخت کرده است.
- ▶ چالش‌های اساسی در ادومتری بصری به طور عمده مربوط به هزینه‌ی پردازشی و شرایط نوری و محیطی است.
- ▶ در این روش وقتی تصاویر دارای بافت کم هستند و یا تلبش شدید و مستقیم وجود دارد، میزان خطا افزایش می‌یابد.
- ▶ وجود سایه‌های ثابت و یا متحرک در صحنه موجب خطا در سیستم می‌شود همچنین سایه خود ربات متحرک می‌تولند موجب خطا در محاسبه جابه‌جایی پیکسل‌ها شود.

کاربردهای Visual Odometry

- ▶ ادومتری بصری کاربردهای گسترده‌ای دارد که به طور موثری در زمینه‌های مختلفی مورد استفاده قرار گرفته است. از کاربردهای آن می‌توان به رباتیک، وسایل نقلیه خودران و وسایل محاسباتی پوشیدنی اشاره کرد. ادومتری بصری در انواع مختلف ربات‌های متحرک مورد استفاده قرار می‌گیرد می‌توان به ربات‌های هوایی، خشکی، روی آب، زیر آب و فضایی اشاره کرد. ناسا در کاوشگر خود در سطح مریخ‌لزد ادومتری بصری برای تعیین مسیر کاوشگر و دوری از مسیرهای پرخطر استفاده می‌کند.
- ▶ ادومتری بصری در ناوبری برای رسیدن به هدف در حالی که از موانع دوری می‌شود مورد استفاده قرار می‌گیرد.
- ▶ در ربات‌های هوایی از ادومتری بصری برای فرود و بلند شدن اتوماتیک استفاده می‌شود.
- ▶ در زمینه‌ی ماشین‌های هوشمند بسیار از ادومتری بصری استفاده می‌شود و بر مبنای آن امکانات هوشمندی ارائه می‌شود مانند ترمز اتوماتیک در زمان‌هایی که ماشین مسیر معقولی ندارند.
- ▶ همچنین در کشاورزی و ماشین‌آلات مربوط به کشاورزی برای تخمین مسیر ماشین‌ها در زمین کشاورزی و دوری از محصولات از ادومتری بصری استفاده می‌شود.

کاربردهای Visual Odometry



▶ در کارخانه‌ها و معادن، از ادومتری بصری در ربات‌ها و ماشین‌هایی که مسیر مشخصی را باید طی کنند استفاده می‌شود.

▶ در **Augmented reality** از ادومتری بصری استفاده می‌شود.



انواع دوربین‌هایی که در Visual Odometry استفاده می‌شوند



Visual Odometry را می‌توان بر اساس دوربینی که استفاده می‌شود دسته‌بندی کرد

انواع مختلفی از دوربین‌ها را می‌توان در ادومتری بصری استفاده کرد. از این دوربین‌ها می‌توان به دوربین‌های Stereo، Monocular، دوربین‌های (Omnidirectional) همه‌جانبه و دوربین‌های RGB-D اشاره کرد.

بیشتر پژوهش‌های حوزمی ادومتری بصری معمولاً از دوربین‌های استریو و مونوکولار استفاده کرده‌اند.

انواع دوربین‌هایی که در Visual Odometry استفاده می‌شوند

Type of VO camera	Pros	Cons
Monocular	Low cost and easy deployment Light weight: good for small robotics Simple calibration	Suffer from image scale uncertainty
Stereo	Image scale and depth information is easy to be retrieved Provide 3D vision	More expensive and needs more calibration effort than monocular cameras It is degraded to the monocular case when the stereo baseline is much smaller than the distances to the scene from the camera Difficult interfacing and synchronization.
Omnidirectional	Provides very wide field of vision (FOV) (up to 360° FOV) Can generate well refined 3D model of the world structure Rotational invariance	Complex system Multiple cameras calibrating and synchronizing Needs high bandwidth Expensive

روش‌های مختلف Visual Odometry

▶ روش‌های تعیین موقعیت ربات‌های متحرک در محیط به وسیله‌ی ادومتری بصری معمولاً به سه صورت انجام می‌پذیرد:

روش‌های مبتنی بر ویژگی: روش مبتنی بر ویژگی شامل سه بخش اصلی استخراج ویژگی از تصویر (گوشه، خط، منحنی و ...)، تطابق یا ردگیری ویژگی‌های استخراج شده و محاسبه‌ی حرکت بر اساس جابه‌جایی پیکسل‌ها است. در این روش‌ها عمل تطابق با مقایسه توصیفگرهای ویژگی‌های استخراج شده انجام می‌پذیرد.

روش‌های مبتنی بر ظاهر: در روش مبتنی بر ظاهر ویژگی از تصاویر استخراج نمی‌شود و از تغییرات شدت روشنایی پیکسل‌ها استفاده می‌شود.

روش‌های ترکیبی: روش‌های ترکیبی هم زمان از دو روش مبتنی بر ویژگی و روش مبتنی بر ظاهر استفاده می‌کند.

Monocular Visual Odometry

اساس و پایه روش

Visual Odometry

Part I: The First 30 Years and Fundamentals

By Davide Scaramuzza and Friedrich Fraundorfer



**University of
Zurich** ^{UZH}

**Zurich Open Repository and
Archive**

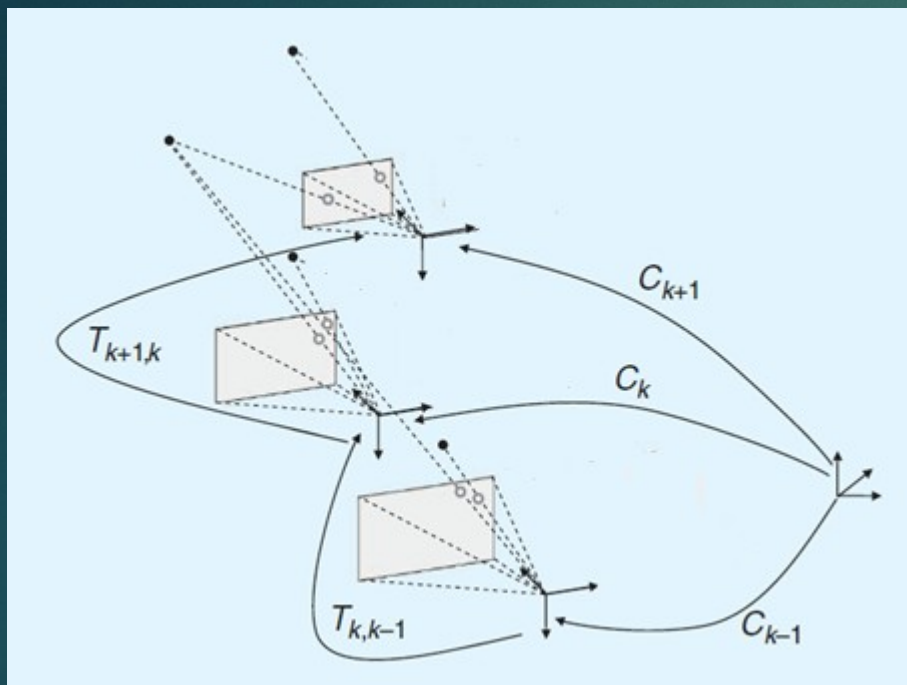
University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2012

Visual Odometry: Part II - Matching, Robustness, and Applications

Fraundorfer, Friedrich; Scaramuzza, Davide

فرمول بندی مسئله Monocular Visual Odometry



یک عامل در یک محیط حرکت می کند و بایا استفاده از
 دوربین متصل به خود هر k ثانیه یک تصویر از محیط
 می گیرد. در روش مونوکولار چون در هر بار فقط یک
 تصویر از صحنه گرفته می شود مجموعه ی تصاویر گرفته
 شده از محیط را با استفاده از نشان می دهیم.
 $I_{0:n} = \{I_0, \dots, I_n\}$

موقعیت دوربین در دو فریم متوالی k و $k-1$ ، توسط
 ماتریس $T_{k,k-1}$ مرتبط هستند. ماتریس $T_{k,k-1}$ یک
 انتقال بین با ۳-آل متغیرینا نچرخا تریس اسچوخوش لانتیسی
 جابه جایی ماتریس جابه جایی است.

$$T_{k,k-1} = \begin{bmatrix} R_{k,k-1} & t_{k,k-1} \\ 0 & 1 \end{bmatrix}$$

1

فرمول بندی مسئله Monocular Visual Odometry

- مجموعه $\{C_0, T_1, T_2, \dots, T_k\}$ حرکات T_k های C_0 معلوم است. T_k متعلق به تمام حرکات است. عامل متحرک در محیط است.
- مجموعه $\{C_0, T_1, T_2, \dots, T_k\}$ نسبت به C_0 در مکان C_0 حرکت دور بین در طول C_0 حرکت دور بین به C_0 نسبت به C_0 است. $K=0$ است.
- محل فعلی دور بین با استفاده از محل قبلی دور بین توسط رابطه زیر محاسبه می شود:

$$C_k = C_{k-1} T_{k:k-1}$$

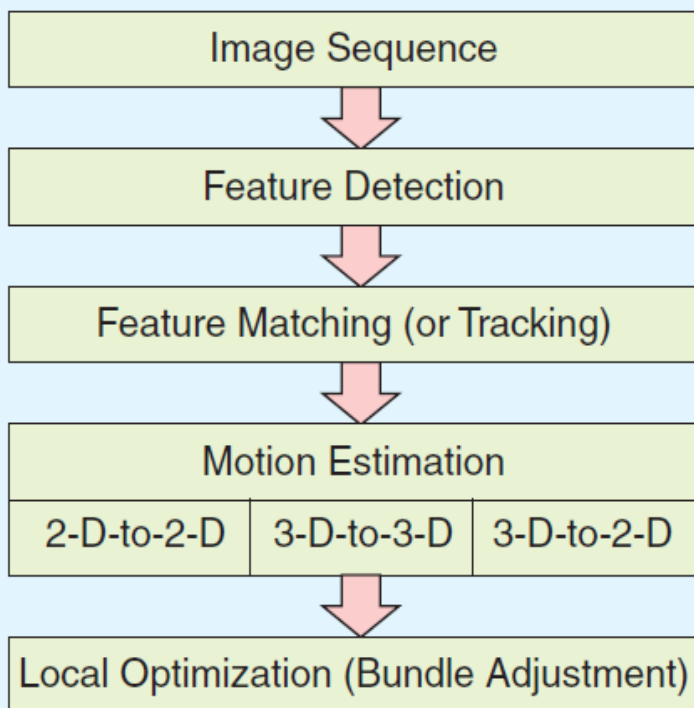
2

- در نتیجه مکان دور بین نسبت به مکان اولیه دور بین C_0 به صورت زیر محاسبه می شود:

$$C_k = C_0 T_{1:0} \dots T_{k:k-1}$$

3

فرمول بندی مسئله Monocular Visual Odometry



➤ **هدف اصلی** در ادومتری بصری محاسبه $T_{k:k-1}$ با استفاده از تصویرهای k و $k-1$ است برای برآوردن پیدنه دقیق به قدری می توان هی توان قدم آخر را مخابر مرتب کردی بهبود و دقت را به نتایج بهتری رسید.

➤ همین طور که قبلا اشاره شد، در روش ادومتری بصری دو روش برای پیدا کردن وجود I_k وجود دارد. روش مبتنی بر ظاهر (زویس پتراسری) وجود دارد. روش مبتنی بر شکلها استفاده می کنند و روش مبتنی بر ویژگی ها استفاده می کنند و روش مبتنی بر ویژگی ها از خصوصیات رنگی هلیتی مانده ظاهر استفاده می کنند. هلیتی که بر روی ظاهر ظاهر می شود و روش مبتنی بر ویژگی ها از خصوصیات رنگی هلیتی مانده ظاهر استفاده می کنند. هلیتی که بر روی ظاهر ظاهر می شود و روش مبتنی بر ویژگی ها از خصوصیات رنگی هلیتی مانده ظاهر استفاده می کنند. هلیتی که بر روی ظاهر ظاهر می شود و روش مبتنی بر ویژگی ها از خصوصیات رنگی هلیتی مانده ظاهر استفاده می کنند.

➤ در شکل روبه روی نمایشی کلی از ادومتری بصری کوپلر کوپلر مبتنی بر استخراج ویژگی را مشاهده می کنید:

محاسبه حرکت

$$E_k \simeq \hat{t}_k R_k,$$

where $t_k = [t_x, t_y, t_z]^T$ and

$$\hat{t}_k = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}.$$

رابطه‌های هندسی بین دو تصویر I_k و I_{k-1} در یک دوربین دو بین کالیبره شده
شده توسط Essential Matrix مشخص می‌شود. ملاترینس
المانی را با E نشان می‌دهند.

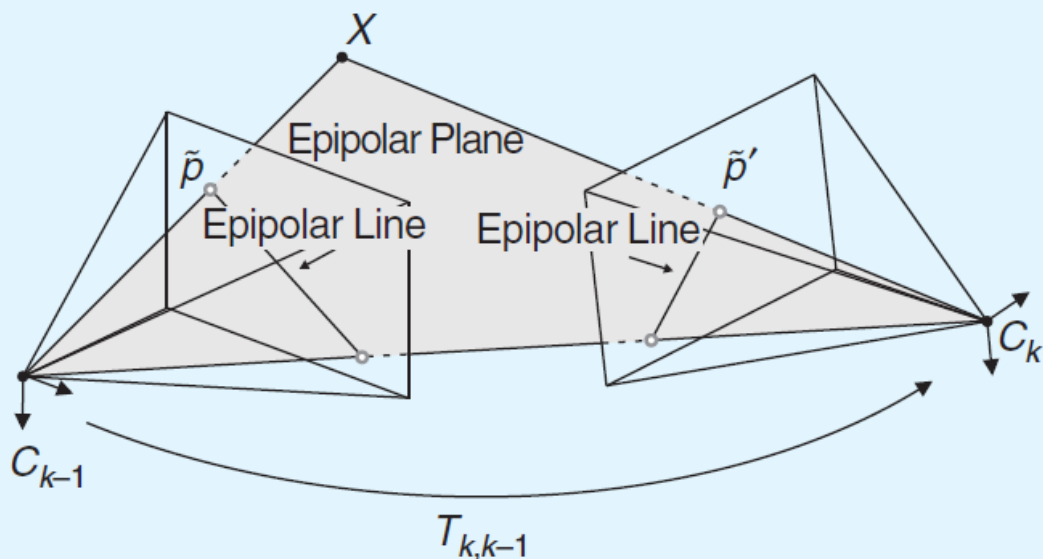
می‌توان ماژیتیس E را بر اساس سیمبل مختل و گیرلی گیرلی و می‌توانی که مختل
مختل محاسبه کرد و سپس ماتریس مختل و مختل جایی را از آن
استخراج کرد.

فرض کنید که \tilde{P} و تصویر یک نقطه در فضای سه بعدی بعدی
در دو زمان مختلف k و $k-1$ باشند. مختصات مختل این دو نقطه را به
صورت زیر نمایش داده می‌شود:

$$\tilde{P}' = [\tilde{u}', \tilde{v}', 1]$$

$$\tilde{P} = [\tilde{u}, \tilde{v}, 1]$$

در شکل روبه‌رو دو نقطه \tilde{P} و \tilde{P}' که تصویرهای یک نقطه یکسان در
فضای سه بعدی و در دو زمان متفاوت هستند، نمایش داده شده
است.



مهم‌تری ویژگی سیستم اپیپولار محدودیت رابطه (۵) است:

$$\tilde{\mathbf{p}}' \mathbf{E} \tilde{\mathbf{p}} = 0$$

5

معمول‌ترین روشی که برای محاسبه‌ی ماتریس \mathbf{E} استفاده می‌شود، روش **Nister's five-point algorithm** است. روش نیستر مکانیزمی دارد که داده‌های پرت را شناسایی می‌کند و در مقابل داده‌های پرت تا حد مناسبی مقاوم است. این الگوریتم فقط از ۵ جفت نقطه متناظر در دو تصویر استفاده می‌کند. روش ۵-نقطه به عنوان روش استاندارد برای پیدا کردن \mathbf{E} شناخته می‌شود. ولی ساده‌ترین روش برای محاسبه‌ی \mathbf{E} ، استفاده از روش هشت-نقطه است. در روش هشت نقطه برای هر دو نقطه متناظر فرمول (۵) به شکل زیر نوشته می‌شود:

$$[\tilde{u}\tilde{u}' \quad \tilde{u}'\tilde{v} \quad \tilde{u}' \quad \tilde{u}\tilde{v}' \quad \tilde{v}\tilde{v}' \quad \tilde{v}' \quad \tilde{u} \quad \tilde{v} \quad 1] \mathbf{E} = 0,$$

6

محاسبه حرکت

▶ با روی هم چیدن دستگاه برای **هشت جفت نقطه**، دستگاه زیر به دست می‌آید:

$$AE = 0$$

7

▶ با حل دستگاه (۷)، پارامترهای E به دست می‌آید. این معادله به راحتی توسط روش تجزیه ماتریس **اس.وی.دی** قابل حل کردن است.

▶ بعد از به دست آوردن ماتریس E ، می‌توان R و t را از آن استخراج کرد. به صورتی عمومی **چهار جواب** مختلف برای R و t به ازای یک ماتریس E وجود دارد. با استفاده از مثلث‌بندی برای یک نقطه می‌توان از میان چهار جواب ممکن، جواب اصلی را پیدا کرد. چهار دسته جلوب به صورت زیر به دست می‌آید:

$$R = U(\pm W^T)V^T,$$

$$\hat{t} = U(\pm W)SU^T,$$

8

در فرمول‌های (۸)، W مقداری برابر با عبارت زیر دلرد: 

$$W^T = \begin{bmatrix} 0 & \pm 1 & 0 \\ \mp 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

9

در شکل زیر، خلاصه‌ی چیزی که گفته شد آورده شده است. 

- 1) Capture new frame I_k
- 2) Extract and match features between I_{k-1} and I_k
- 3) Compute essential matrix for image pair I_{k-1}, I_k
- 4) Decompose essential matrix into R_k and t_k , and form T_k
- 5) Compute relative scale and rescale t_k accordingly
- 6) Concatenate transformation by computing $C_k = C_{k-1} T_k$
- 7) Repeat from 1).

انتخاب ویژگی و پیدا کردن نقاط متناظر

- ▶ دو روش کلی برای پیدا کردن نقاط متناظر وجود دارد: ۱-ردگیری ۲-تطبیق.
- ▶ در روش ردگیری یک سری نقطه در تصویر لانتخاب می‌شود و این نقاط در تصویرهای بعدی ردگیری می‌شوند. در این روش‌ها معمولاً از روش‌های مبتنی بر شارنوری و مدل‌سازی‌های حرکت (مانند کالمن) برای ردگیری استفاده می‌شود.
- ▶ در روش تطبیق از دو تصویر متوالی یک سری ویژگی‌ها به عنوان نقطه‌ی کلیدی انتخاب می‌شوند و برای هر کدام توصیفگری ایجاد می‌شود. سپس بر اساس فاصله‌ی توصیفگر نقطه‌های کلیدی، نقاط متناظر پیدا می‌شوند.
- ▶ روش اول برای زملنی که دو تصویر در فاصله زملنی کمی گرفته شده‌اند مناسب است و روش دوم برای زمانی مناسب که است که حرکتی بزرگ دو فریم پشت سر هم رخ داده است.
- ▶ در مرحله‌ی استخراج ویژگی، تصویر برای پیدا کردن نقاطی جست‌وجو می‌شود. این نقاط باید به گونه‌ای باشند که در سایر تصویرهای بعدی هم بتوان آن‌ها را پیدا کرد و عمل تطبیق را انجام داد. برای عمل ادومتری بصری، عموماً دو نوع نقطه کلیدی گوشه (Corner) و یا حباب (Blob) استفاده می‌شود.

انتخاب ویژگی و پیدا کردن نقاط متناظر

- ▶ گوشه (Corner) به محل تلاقی دو یا بیش از دو لبه می‌گویند.
- ▶ یک حباب در تصویر از همسایه‌هایش از نظر بافت، رنگ و شدت روشنایی متفاوت است ولی یک گوشه و لبه نیست.
- ▶ ویژگی‌ها باید یک سری خصوصیت‌های اساسی داشته باشند: ویژگی‌ها باید قابلیت موقیت‌یابی دقیق در تصویر را داشته باشند، ویژگی‌های باید در مقیاس‌های مختلف استخراج شوند، ویژگی‌ها باید تکرارپذیر باشند یعنی بتوان تعداد زیادی از آن‌ها را در تصویر بعدی پیدا کرد، باید از نظر زمینی و محاسباتی به صرفه باشند، ویژگی‌ها باید در مقابل تغییرات هندسی مانند جابه‌جایی، چرخش و تغییر مقیاس مقاوم باشند.
- ▶ با توجه به روشی که برای ادومتری بصری استفاده می‌کنیم و با توجه به محیطی که عامل در آن قرار دارد، ویژگی و یا ویژگی‌های مناسبی را از میان روش‌های موجود پیدا کنیم.
- ▶ ویژگی‌های گوشه از نظر محاسباتی سریع هستند ولی تفکیک‌پذیری کمتری دارند. ویژگی‌های حباب تفکیک‌پذیرتر هستند ولی محاسبات سنگین‌تری دارند. گوشه‌ها به خوبی موقعیتشان در تصویر مشخص می‌شود ولی در فضای مقیاس سخت‌تر پیدا می‌شوند.

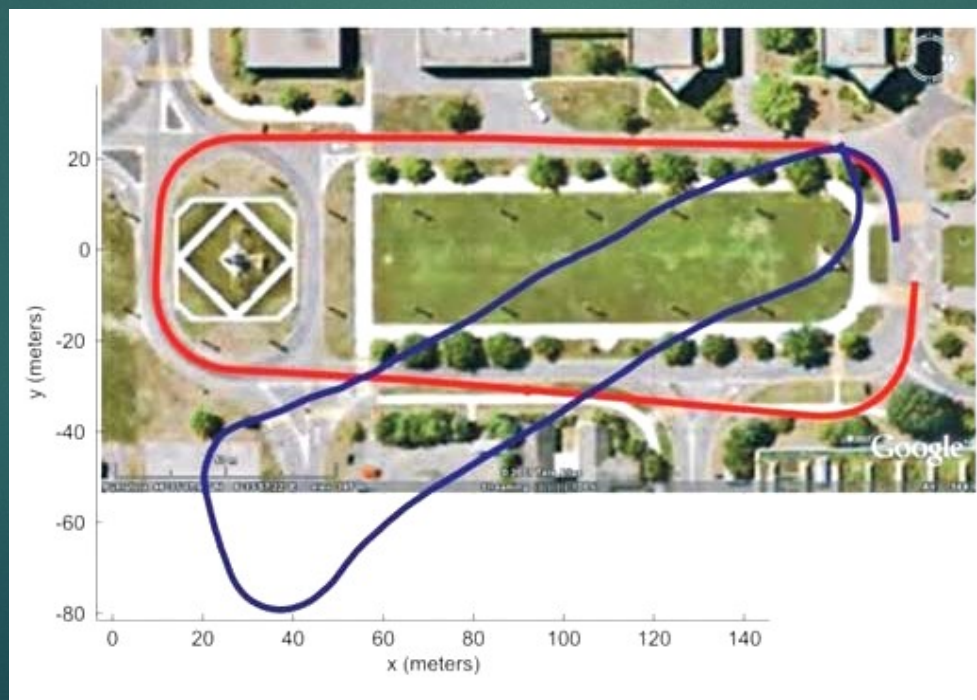
انتخاب ویژگی و پیدا کردن نقاط متناظر

► برای ادومتری بصری براساس محیط، قدرت محاسباتی، برخط بودن یا نبودن و...، باید ویژگی و الگوریتم مناسبی برای انتخاب ویژگی انتخاب کرد.

	Corner detector	Blob detector	Rotation invariant	Scale invariant	Affine invariant	Repeatability	Localization accuracy	Robustness	Efficiency
Harris	x		x			+++	+++	++	++
Shi-Tomasi	x		x			+++	+++	++	++
FAST	x		x	x		++	++	++	++++
SIFT		x	x	x	x	+++	++	+++	+
SURF		x	x	x	x	+++	++	++	++
CenSurE		x	x	x	x	+++	++	+++	+++

حذف داده‌های پرت

معمولا نقاط متناظری که پیدا می‌شوند حاوی داده‌های پرت (Outlier) هستند. علت اصلی وجود داده‌های پرت نویز، هم‌پوشانی، بلورشدگی و تغییر در زلویی دید و نورپردازی است. بسیاری از روش‌های استخراج ویژگی مدلی ساده برای تغییر، نورپردازی و چرخش دوربین در نظر می‌گیرند و راهی برای مقابله با داده‌های پرت ارائه نمی‌دهند. داده‌های پرت به شدت عملکرد را تحت تاثیر قرار می‌دهند.



حذف داده‌های پرت

یکی از مشهورترین و ساده‌ترین روش‌ها برای حذف داده‌های پرت روش (RANSAC) رنسک است. ایده‌ی اصلی این روش نمونه‌برداری تصادفی است. در این روش به صورت تصادفی نمونه‌برداری صورت می‌پذیرد و با نمونه‌های انتخاب شده ماتریس R و t محاسبه می‌شوند. بعد با استفاده از داده‌هایی که انتخاب نشده‌اند امتیاز سیستم حساب می‌شود. این کار بارها تکرار می‌شود و در آخر بهترین مدل انتخاب می‌شود. بر اساس فاصله نقاط تا مدل نهایی، نقاط به دو دسته داده‌های پرت و داده‌های غیرپرت تقسیم می‌شوند.

Algorithm 1: RANSAC

- 1) Initial: let A be a set of N feature correspondences
- 2) repeat
 - 2.1) Randomly select a sample of s points from A
 - 2.2) Fit a model to these points
 - 2.3) Compute the distance of all other points to this model
 - 2.4) Construct the inlier set (i.e. count the number of points whose distance from the model $< d$)
 - 2.5) Store these inliers
 - 2.6) until maximum number of iterations reached
- 3) The set with the maximum number of inliers is chosen as a solution to the problem
- 4) Estimate the model using all the inliers

Monocular Visual Odometry برخی از روش‌های

2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics

A Classification-Based Visual Odometry Approach

Wang Zhou, Hao Fu, Xiangjing An
*College of Mechatronic Engineering and Automation
 National University of Defense Technology
 Changsha, P.R.China, 410073
 bahai_nudt@163.com*

DeepVO: A Deep Learning approach for Monocular Visual Odometry

Vikram Mohanty Shubh Agrawal Shaswat Datta Arna Ghosh
 Vishnu D. Sharma Debashish Chakravarty
 Indian Institute of Technology Kharagpur
 Kharagpur, West Bengal, India 721302

{vikram.mohanty, shubh.agrawal111, shaswatdatta, arna.ghosh, vds, dc}@iitkgp.ac.in

Monocular Visual Odometry Based on Optical Flow and Feature Matching

Cheng Chuanqi^{1,2}, Hao Xiangyang^{1,2}, Zhang Zhenjie^{1,2}, Zhao Mandan^{1,2}

1. School of Navigation and Aerospace Engineering, Information Engineering University, Zhengzhou 450001, China

2. Beidou Navigation Technology Collaborative Innovation Center of Henan, Zhengzhou 450001, China

E-mail: legend3q@163.com xiangyanghao2004@163.com zzxiaodao@126.com mandanzhao@163.com

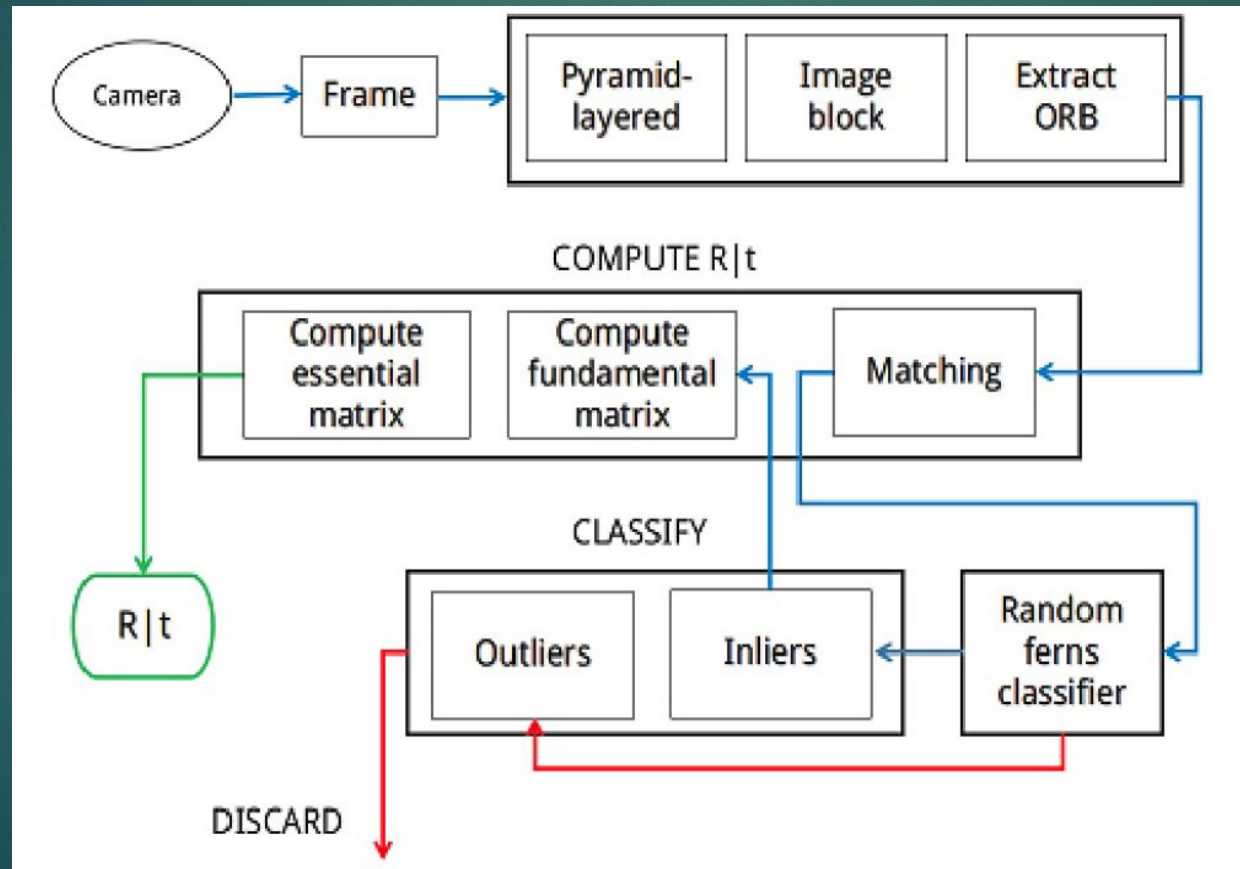
Abstract: To solve the problem of real-time precise localization in CPS-denied places, a monocular visual odometry method based on optical flow tracking and feature matching is proposed. To speed up traditional pose estimation algorithm, the image sequences are classified into key frames and non-key frames. The conventional pipeline of feature detection and matching is utilized to process key frames, while utilizing Lucas Kanade optical flow to track the correspondences in non-key frames. To improve the robustness of the visual odometry method, a RANSAC-based outlier rejection scheme is applied in the phase of pose estimation. Moreover, a Kalman Filter based on the dynamic

A Classification-Based Visual Odometry Approach

- ▶ در این مقاله، یک روش مبتنی بر دسته‌بندی برای ادومتری بصری مونوکولار پیشنهاد شده است.
- ▶ در این روش دقیقاً مانند چیزی که تا کنون گفته شد عمل می‌شود.
- ▶ در این روش از ویژگی ORB استفاده می‌شود. برای پیدا کردن R و t از روش ۵-نقطه استفاده می‌کند.
- ▶ برای حذف داده‌های پرت از روش رنسک بهره می‌برد.
- ▶ در روش ارائه شده، ویژگی‌هایی که توسط رنسک به عنوان دادمی پرت شناسایی شده‌اند دریافت می‌شوند و سپس با استفاده از داده‌های پرت و داده‌های غیرپرت و توصیفگرهایشان یک دسته‌بند باینری آموزش داده می‌شود.
- ▶ این دسته‌بند به صورت برخط (Online) آموزش داده می‌شود.
- ▶ بعد از گذشت چند فریم بخش دسته‌بند فعال می‌شود. بعد از استخراج نقاط کلیدی و ایجاد توصیفگرهایشان، نقاط کلیدی را به دسته‌بند می‌دهیم. دسته‌بند اعلام می‌کند که کدام نقطه کلیدی مناسب است و کدام مناسب نیست. نقاط بد را که دسته‌بند تشخیص می‌دهد را حذف می‌کنیم.

A Classification-Based Visual Odometry Approach

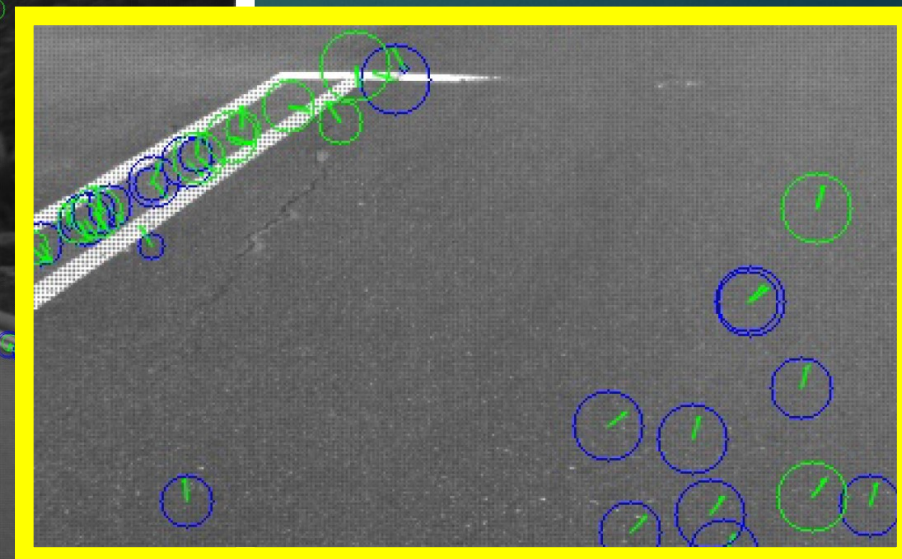
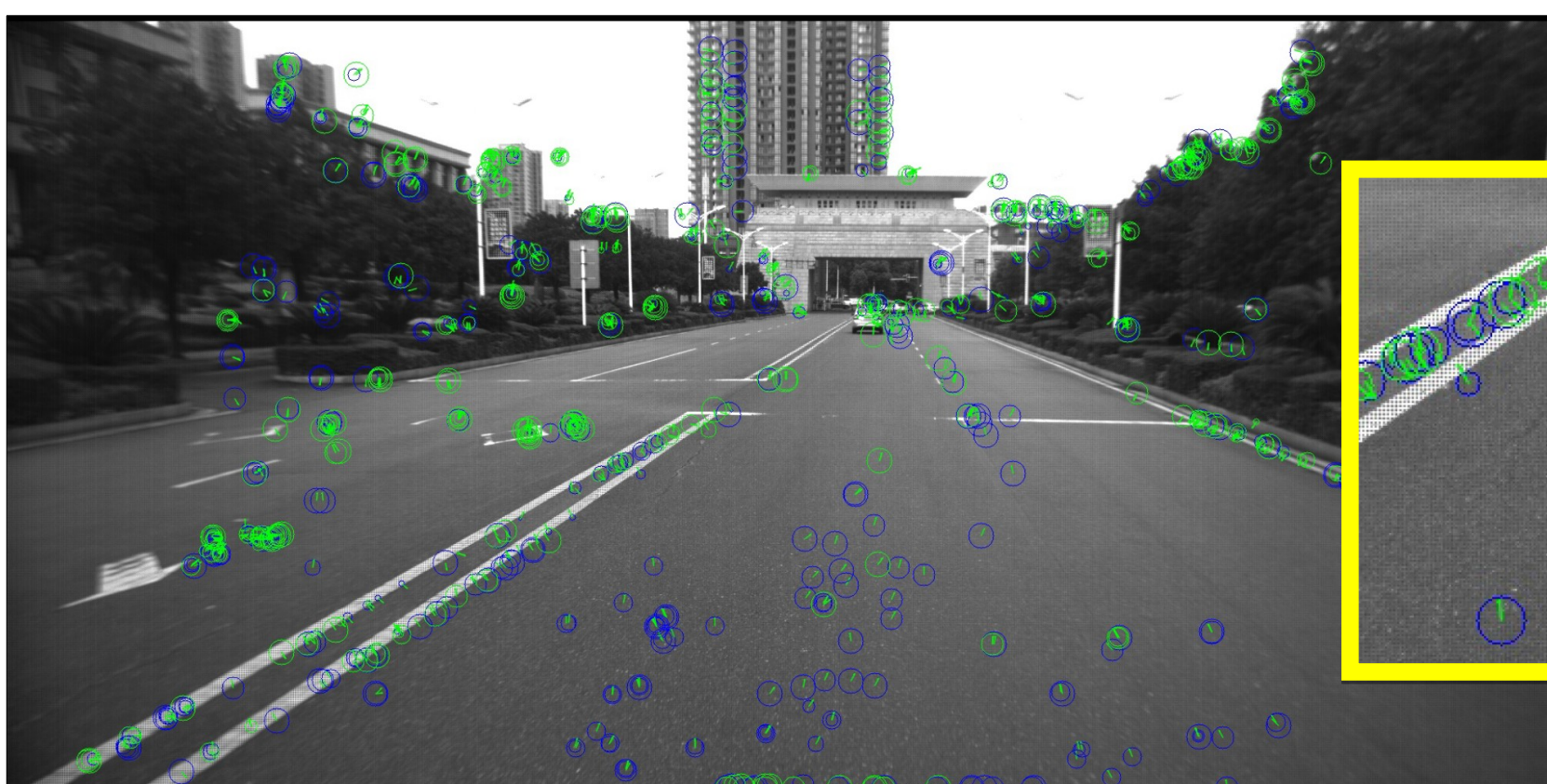
38



A Classification-Based Visual Odometry Approach

39

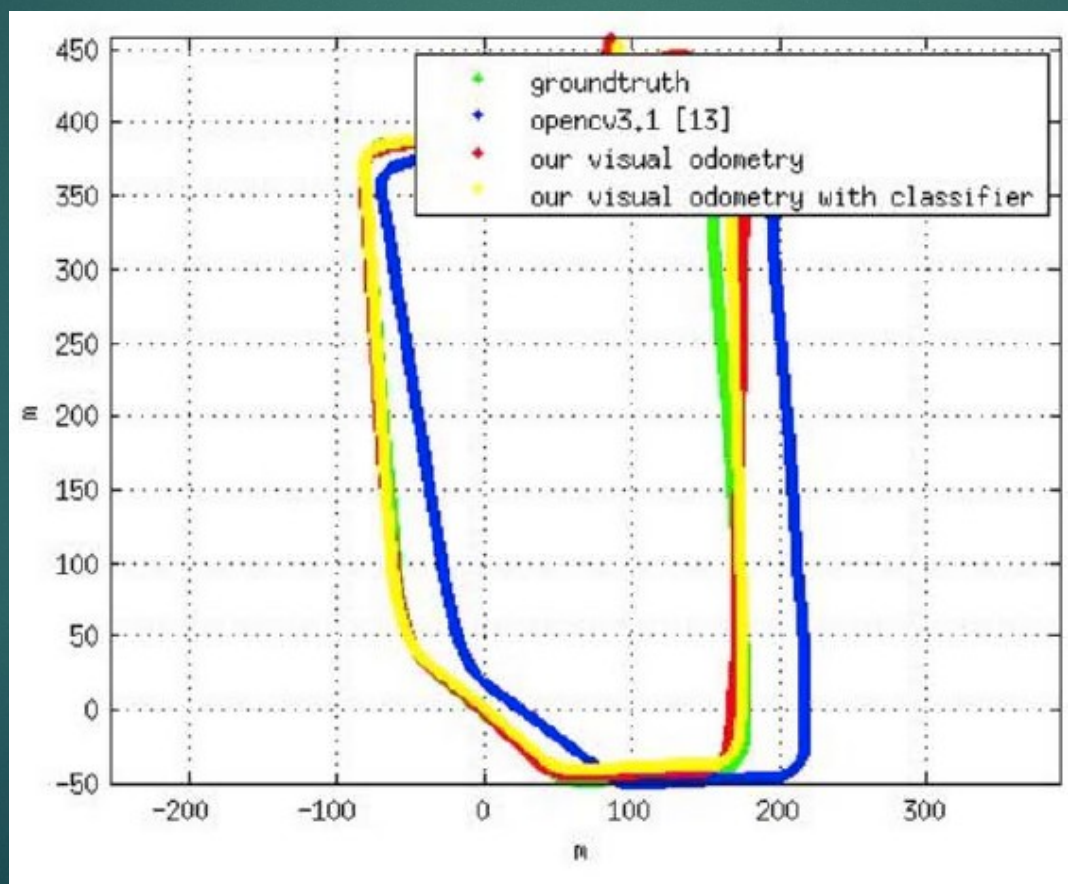
در شکل زیر بعضی از نقاطی که توسط این روش حذف می‌شوند را مشاهده می‌کنید. همین‌طور که مشاهده می‌شود داده‌های یرت در میان نقاط کلیدی، شناسایی شده‌اند. در شکل نقاط نامناسب با رنگ آبی نمایش داده شده‌اند.



A Classification-Based Visual Odometry Approach

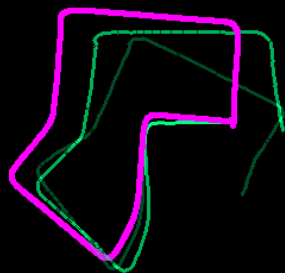
40

تأثیر این روش در خروجی نهایی و مسیر پیدا شده در تصویر زیر قابل مشاهده است. ►

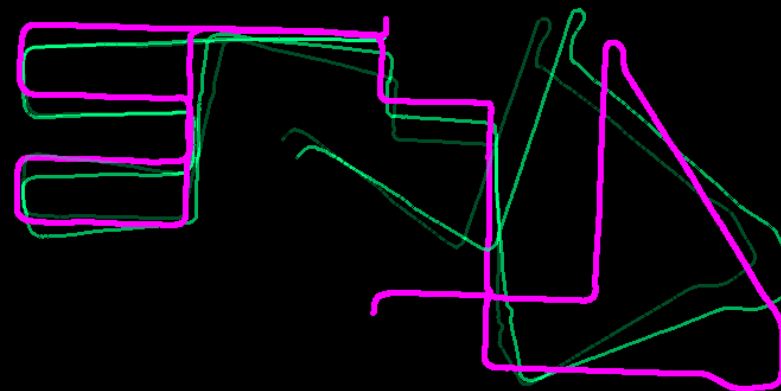


A Classification-Based Visual Odometry Approach

41



Purple: Ground Truth - Green: Output - seq 07 - scale of drawing 1.0



Purple: Ground Truth - Green: Output - seq 08 - scale of drawing 0.8

DeepVO: A Deep Learning approach for Monocular Visual Odometry

- ▶ در این مقاله، به جای استفاده از استخراج و تطبیق ویژگی‌های شبکه‌های عصبی عمیق استفاده شده است که عمل رگرسیون را انجام می‌دهد.
- ▶ در ابتدا دیتاست پیش‌پردازش شده است و دیتاست به گونه‌ای تغییر داده شده است که هر عضو آن به صورت زیر است:

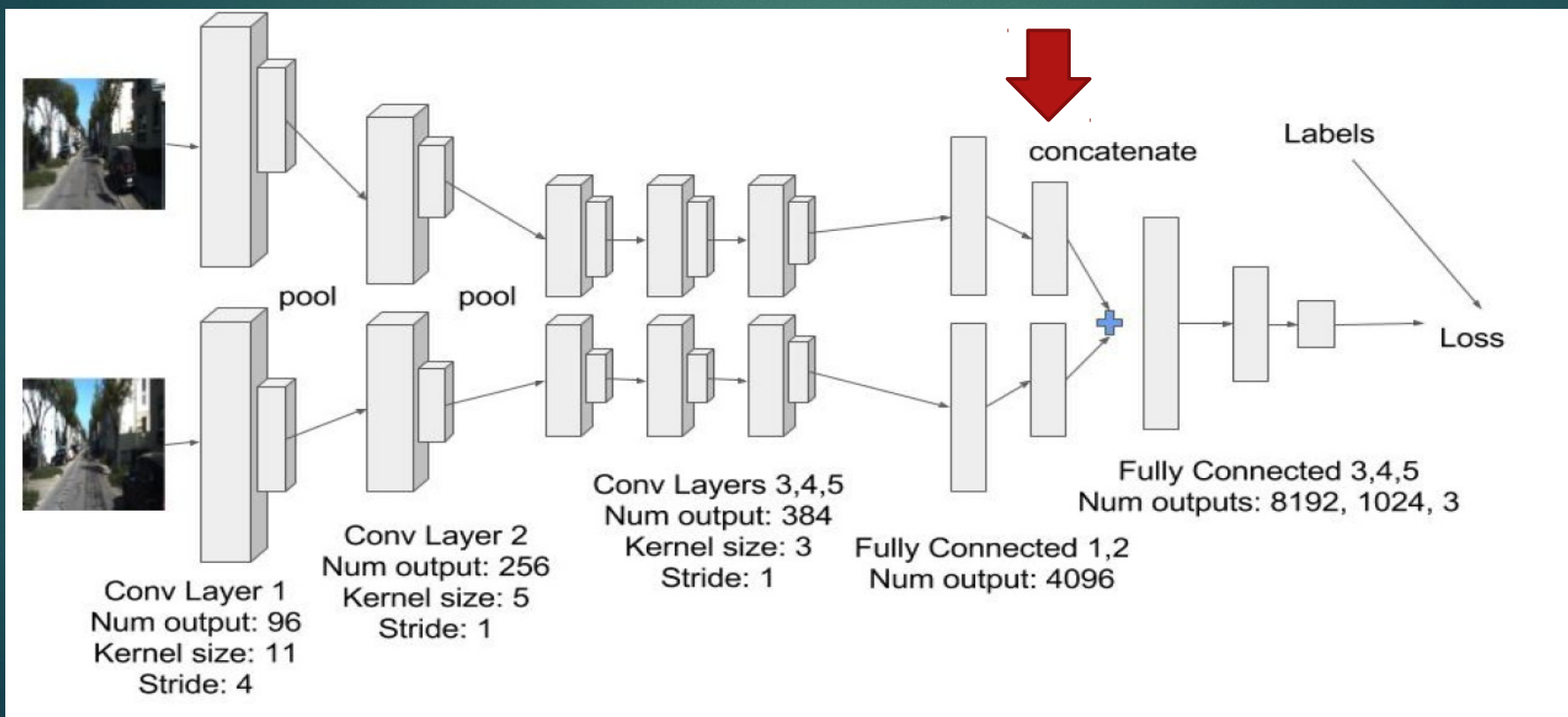
$$\{ I_t, I_{t+1}, (\Delta x, \Delta z, \Delta \Theta)_{t \rightarrow (t+1)} \}$$

- ▶ I_t و I_{t+1} دو فریم متوالی هستند و سایر موارد میزبان جابه‌جایی صورت گرفته. در دو فریم متوالی است
- ▶ برای این کار شبکه عصبی عمیقی مبتنی بر Alexnet استفاده شده است.

DeepVO: A Deep Learning approach for Monocular Visual Odometry

43

در شکل زیر شبکه استفاده شده را مشاهده می کنید. همین طور که در شکل دیده می شود از دو شبکه الکسنت استفاده شده است. دو فریم متوالی به عنوان ورودی به شبکه داده می شود.



DeepVO: A Deep Learning approach for Monocular Visual Odometry

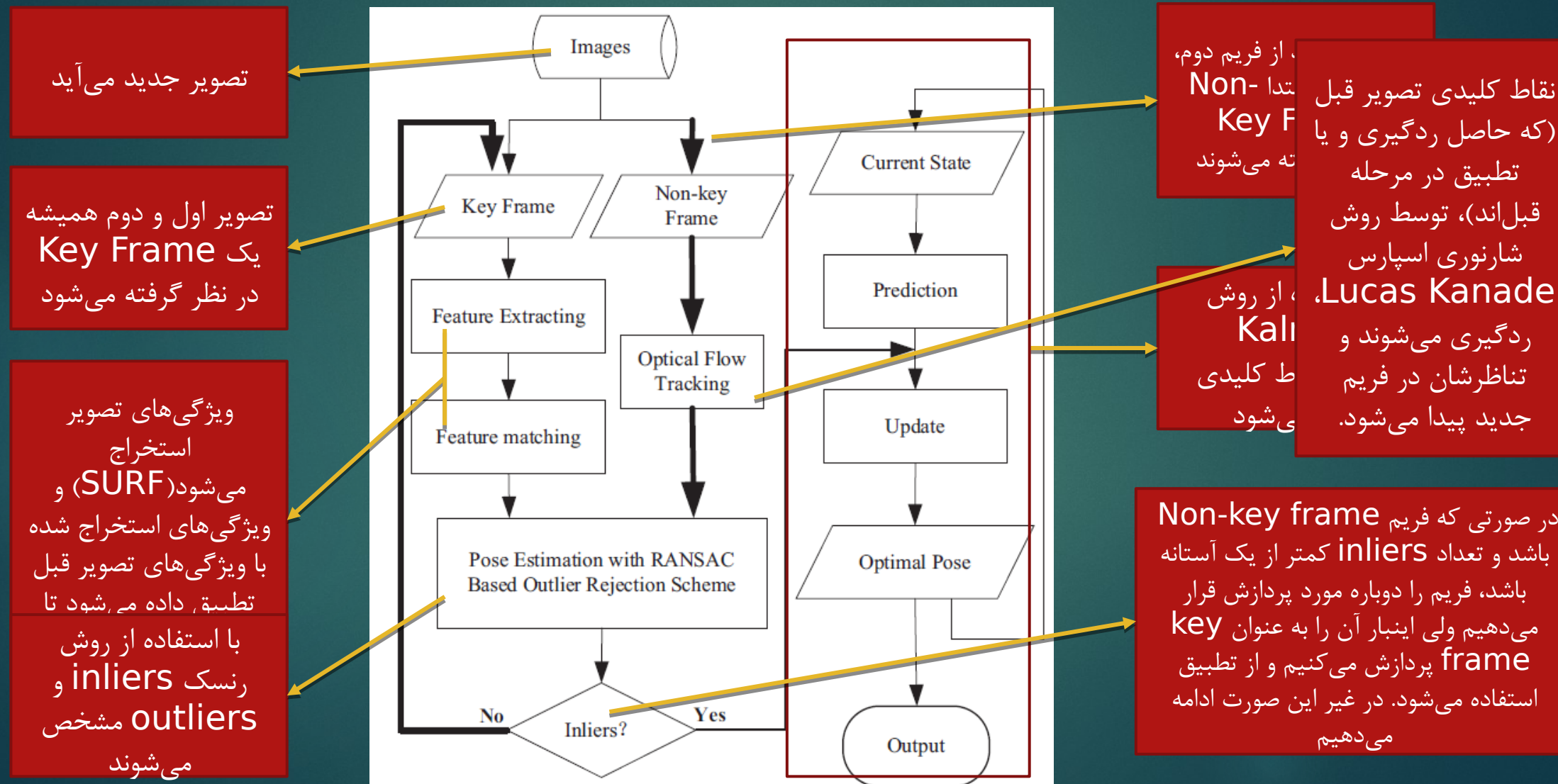
- ▶ سناریوهای مختلفی برای آموزش و تست این شبکه استفاده شده است که شامل موارد زیر است:
- ▶ داده‌های آموزش و تست از ویدیوهای جدا از هم انتخاب شده‌اند. این آزمایش برای بررسی عملکرد شبکه در هنگامی که شبکه اطلاعات قبلی از محیط ندارد است.
- ▶ داده‌های آموزش و تست از ویدیوهای یکسانی برداشته شده‌اند. این آزمایش برای بررسی عملکرد شبکه در هنگامی است که اطلاعاتی از محیط در اختیار شبکه قرار بگیرد.
- ▶ پیش آموزش شبکه با دیتاست **ImageNet**، این آزمایش برای بررسی عمل تشخیص اشیا و تاثیرش بر اودومتري بصری انجام شده است.
- ▶ این روش فقط در زمانی که داده‌های آموزش و تست از مجموعه ویدیوهای یکسانی انتخاب شدند عملکرد خوبی دارد. در سایر موارد عملکرد بسیاری ضعیفی از خود نشان می‌دهد.

Monocular Visual Odometry Based on Optical Flow and Feature Matching

- ▶ در این روش در کنار عمل تطبیق ویژگی‌های کلیدی استخراج شده، در بسیاری از موارد از ردگیری (Tracking) نیز استفاده می‌شود.
- ▶ عمل ردگیری با استفاده از روش **Lucas Kanade optical flow** انجام می‌شود که نوعی شار نوری به صورت **Sparse** است.
- ▶ ردگیری نسبت به عمل تطبیق بسیار سریع‌تر است به همین دلیل به عملکردی بهتر از **Real-Time** می‌رسیم.
- ▶ در ادامه الگوریتم این روش شرح داده شده است.

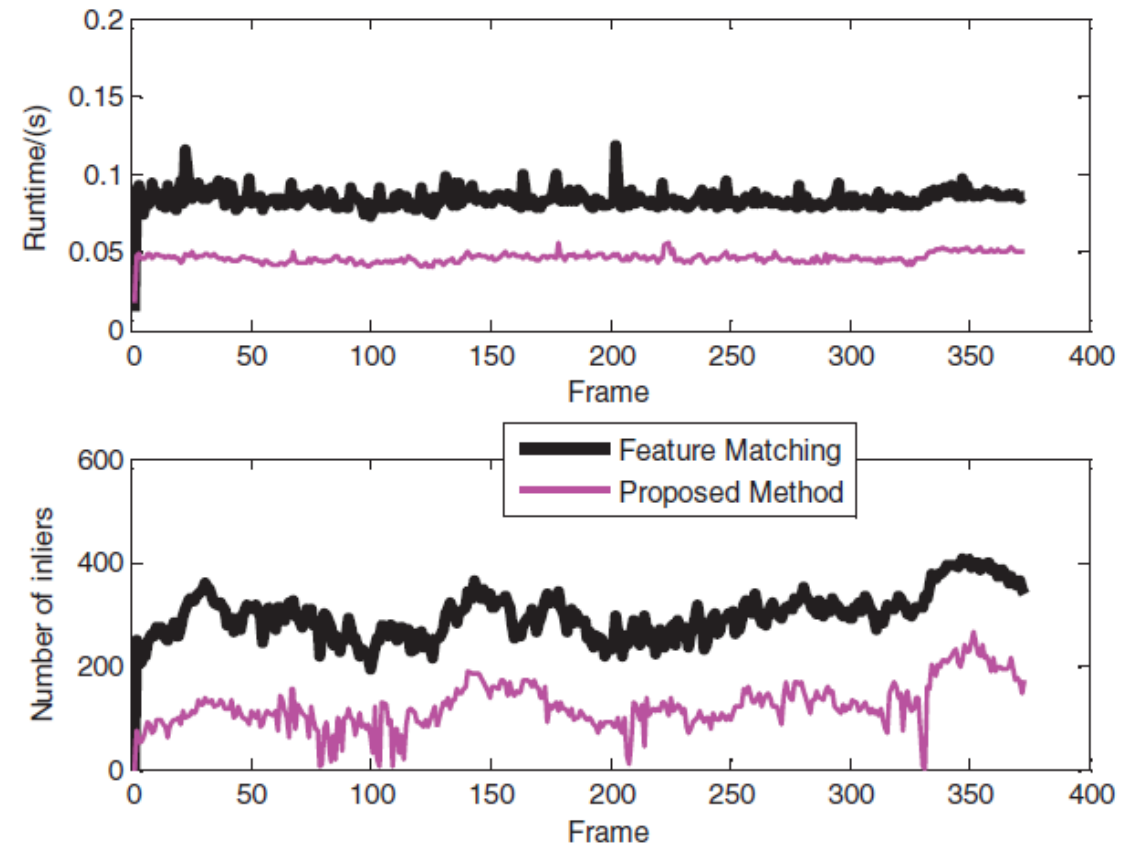
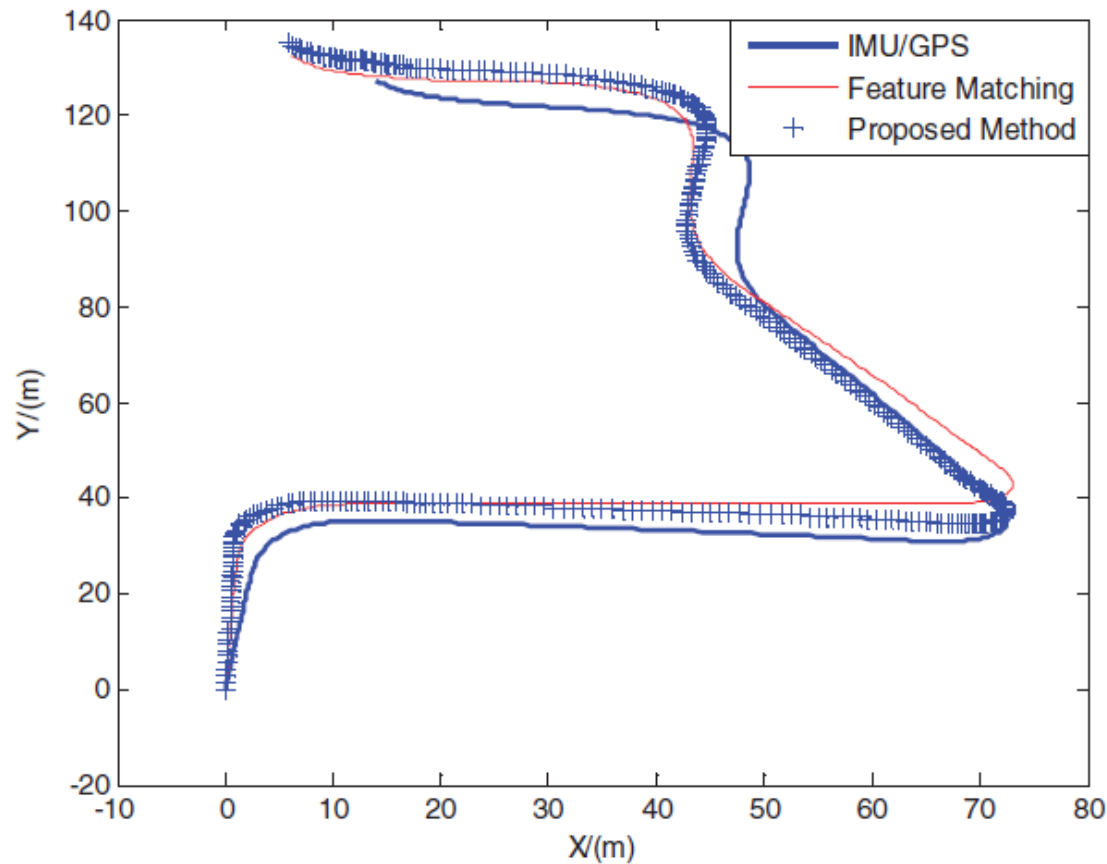
Monocular Visual Odometry Based on Optical Flow and Feature Matching

46



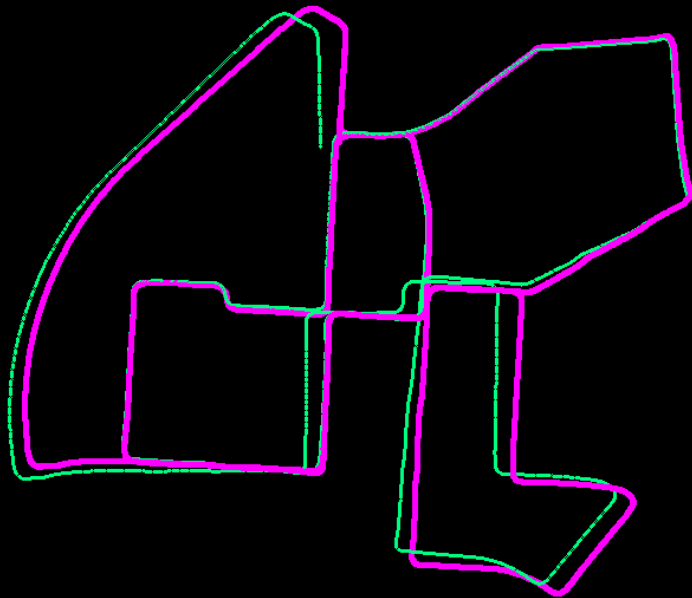
Monocular Visual Odometry Based on Optical Flow and Feature Matching

47



Monocular Visual Odometry Based on Optical Flow and Feature Matching

48



Purple: Ground Truth – Green: Output – seq 00 – scale of drawing 1.0



Purple: Ground Truth – Green: Output – seq 09 – scale of drawing 0.8