

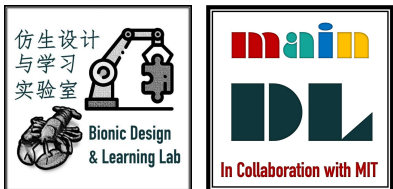
# A Method of Calculating Odometers based on SLAM with Binoplastic Vision

王鑫怡: 12010239 | 黄宝仪: 12010437

张郁洁: 12010142 | 倪健翔: 11812723

邱望弘杰: 12010310

2023.3.6



AncoraSIR.com



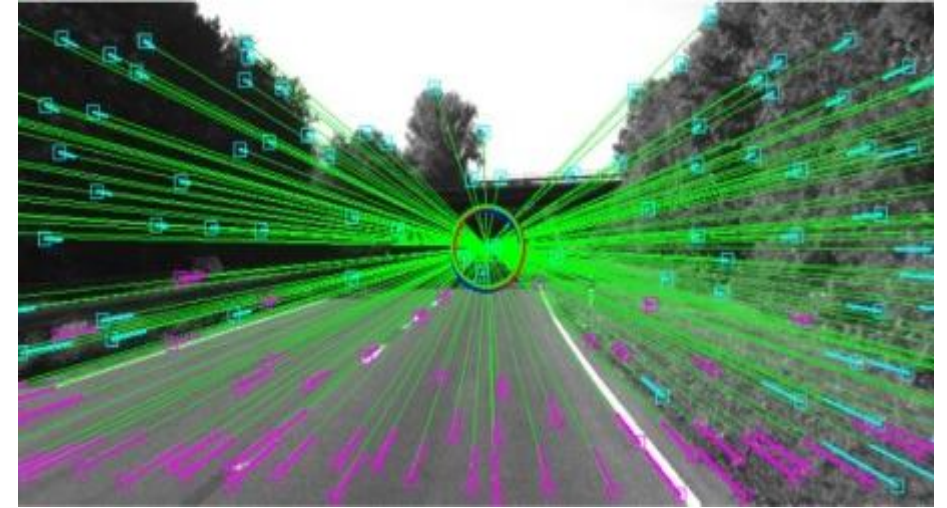
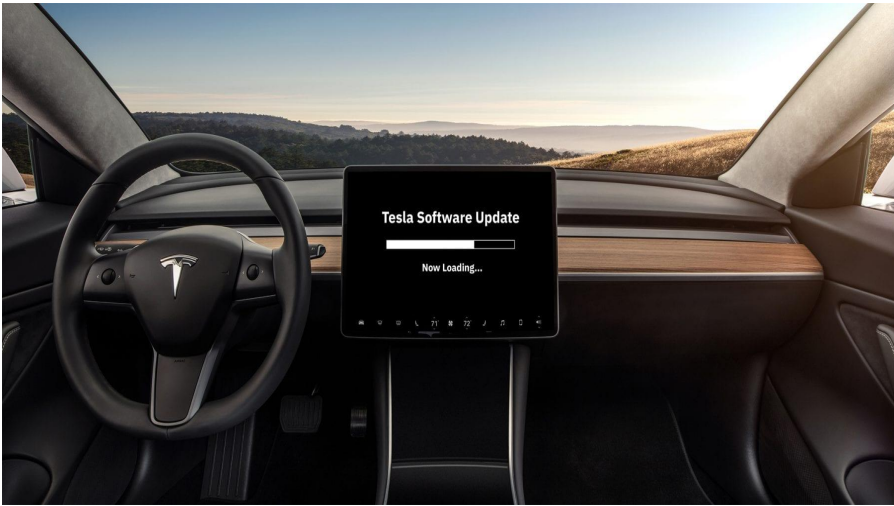
# Summary

---

- The automotive industry is in an era of change, and the development of autonomous driving related technologies has huge market potential. Navigation and positioning is important for automatic driving in cities. Visual positioning is a key part of mobile robots and automatic driving, providing rich information at low cost.
- We reviewed papers to understand algorithms for autonomous navigation:
  - Stereo Visual Odometry for Road Vehicles Based on a Point-to-Epipolar-Line Metric & Legged Locomotion in Challenging Terrains using Egocentric Vision
- We plan to use the KITTI2015 and Middlebury Stereo dataset for model training. The former contains a rich real traffic environment data, while the latter collect mainly indoor environment data.
- We are going to propose currently the highest-ranking algorithm on the KITTI score-board: SOFT2.
- We decided to evaluate our method based on the following variables:
  - Misclassification rate, Precision and recall, Calculate F1, ROC curve, ROC\_AUC, and k-fold cross-validation.

# Problem

## *Navigation of automatic driving based on visual positioning*



**Navigation and positioning** is important for automatic driving in cities.

**Previous method** mostly rely on the **fusion of sensor** such as IMUs and cameras, LIDAR and millimeter-wave radar which each part's data should be precise.

**Visual positioning** is a key part of mobile robots and automatic driving, providing rich information at low cost.

# Reading to Examine

---

- SLAM algorithm
- Public dataset of road vehicle  
e.g. KITTI-360 dataset, Oxford Robotics Car dataset
- Stereo Visual Odometry

- Reference:

Cvisic, Markovic, I., & Petrovic, I. (2022). SOFT2: Stereo Visual Odometry for Road Vehicles Based on a Point-to-Epipolar-Line Metric. *IEEE Transactions on Robotics*, 1–16. <https://doi.org/10.1109/TRO.2022.3188121>

Agarwal, Kumar, A., Malik, J., & Pathak, D. (2022). Legged Locomotion in Challenging Terrains using Egocentric Vision. [arXiv.org](https://arxiv.org/abs/2203.11411).

# Data

- We plan to use the following dataset for model training:
  - KITTI 2015: Real traffic environment for autonomous driving.
  - Middlebury Stereo Datasets: Indoor environment data for stereo matching and depth estimation.

[Stereo](#) [Evaluation](#) [Datasets](#) [Code](#) [Submit](#)

### Middlebury Stereo Datasets

[2001 datasets](#) - 6 datasets of piecewise planar scenes [1]  
(Sawtooth, Venus, Bull, Poster, Barn1, Barn2)

[2003 datasets](#) - 2 datasets with ground truth obtained using structured light [2]  
(Cones, Teddy)

[2005 datasets](#) - 9 datasets obtained using the technique of [2], published in [3, 4]  
(Art, Books, Dolls, Laundry, Moebius, Reindeer, Computer, Drumsticks, Dwarves)

[2006 datasets](#) - 21 datasets obtained using the technique of [2], published in [3, 4]  
(Aloe, Baby1-3, Bowling1-2, Cloth1-4, Flowerpots, Lampshade1-2, Midd1-2, Monopoly, Plastic, Rocks1-2, Wood1-2)

[2014 datasets](#) - 33 datasets obtained using the technique of [5]

[2021 mobile datasets](#) - 24 datasets obtained with a mobile device on a robot arm, using the technique of [5]

AncoraSIR.com

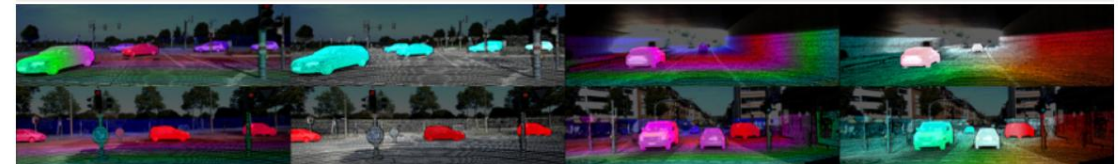
## The KITTI Vision Benchmark Suite

A project of Karlsruhe Institute of Technology and Toyota Technological Institute at Chicago

home setup **stereo** flow sceneflow depth odometry object tracking road semantics raw data submit results

A. Geiger | P. Lenz | C. Stiller | R. Urtasun [Log in](#)

## Stereo Evaluation 2015



The stereo 2015 / flow 2015 / scene flow 2015 benchmark consists of 200 training scenes and 200 test scenes (4 color images per scene, saved in loss less png format). Compared to the stereo 2012 and flow 2012 benchmarks, it comprises dynamic scenes for which the ground truth has been established in a semi-automatic process. Our evaluation server computes the percentage of bad pixels averaged over all ground truth pixels of all 200 test images. For this benchmark, we consider a pixel to be correctly estimated if the disparity or flow end-point error is  $<3\text{px}$  or  $<5\%$  (for scene flow this criterion needs to be fulfilled for both disparity maps and the flow map). We require that all methods use the same parameter set for all test pairs. Our development kit provides details about the data format as well as MATLAB / C++ utility functions for reading and writing disparity maps and flow fields. More details can be found in [Object Scene Flow for Autonomous Vehicles \(CVPR 2015\)](#).

# Method

**SOFT2: currently the highest-ranking algorithm on the KITTI score-board.**

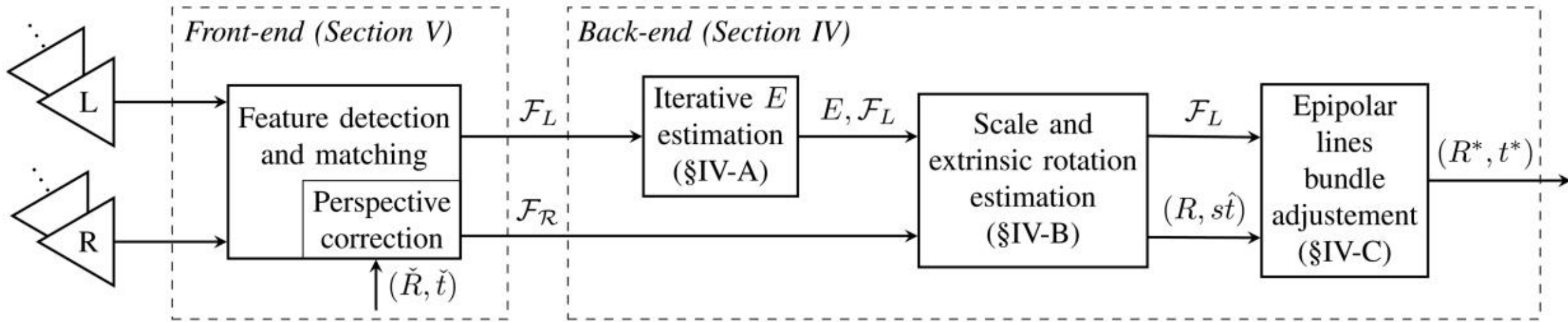


Fig. 2. Pipeline of the proposed stereo odometry approach. Triangles represent left and right cameras (L and R), respectively.

SOFT2 relies on the constraints imposed by the epipolar geometry and kinematics.

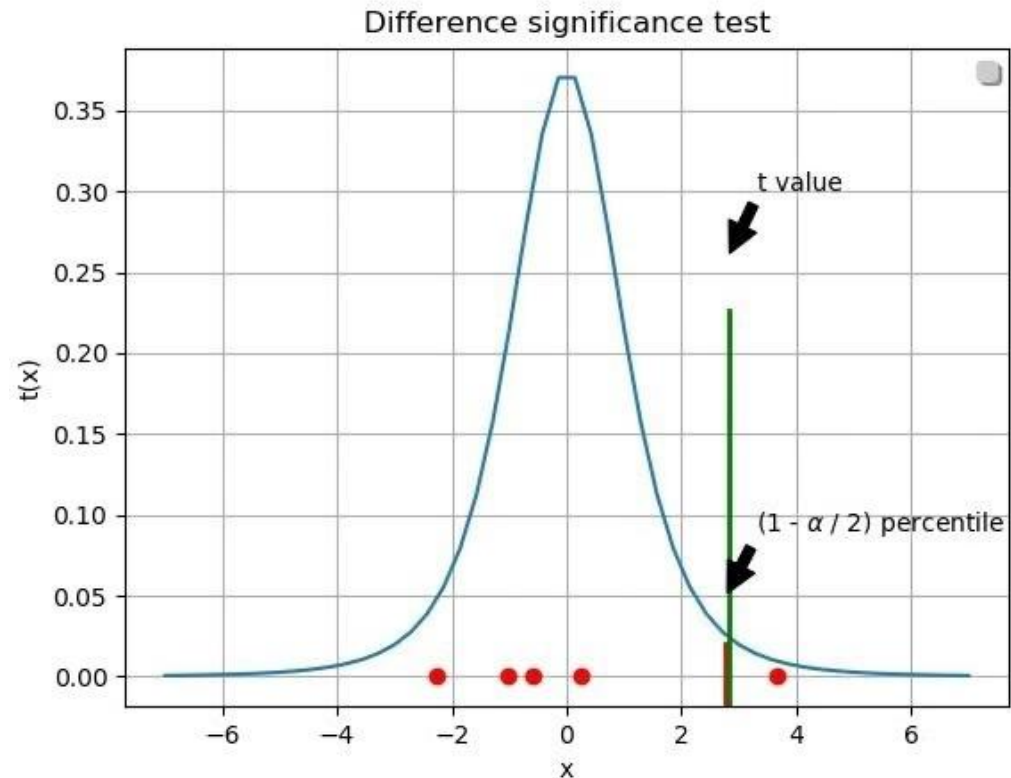
# Method

---

- Minimize point-to-epipolar-line distances.
- Propose to jointly estimate the absolute scale and the extrinsic rotation of the second camera in order to alleviate the effects of varying stereo rig extrinsics.
- Smooth the motion estimates in a temporal window of frames by using the proposed epipolar line bundle adjustment procedure.

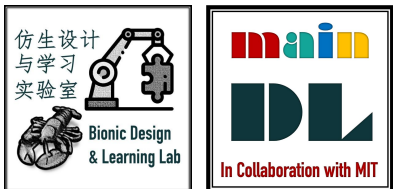
# Evaluation

- We decided to evaluate our method based on the following variables:
  - Misclassification rate (misclassification rate), that is, accuracy.
  - Precision and recall
  - Calculate F1
  - ROC curve, ROC\_AUC
  - k-fold cross-validation





# Thank You



[AncoraSIR.com](http://AncoraSIR.com)

