Capital Normal University

Visual Saliency Oriented Vehicle Scale Estimation

Reporter: Jiali Ding   2020.12.07

Context

PART 1

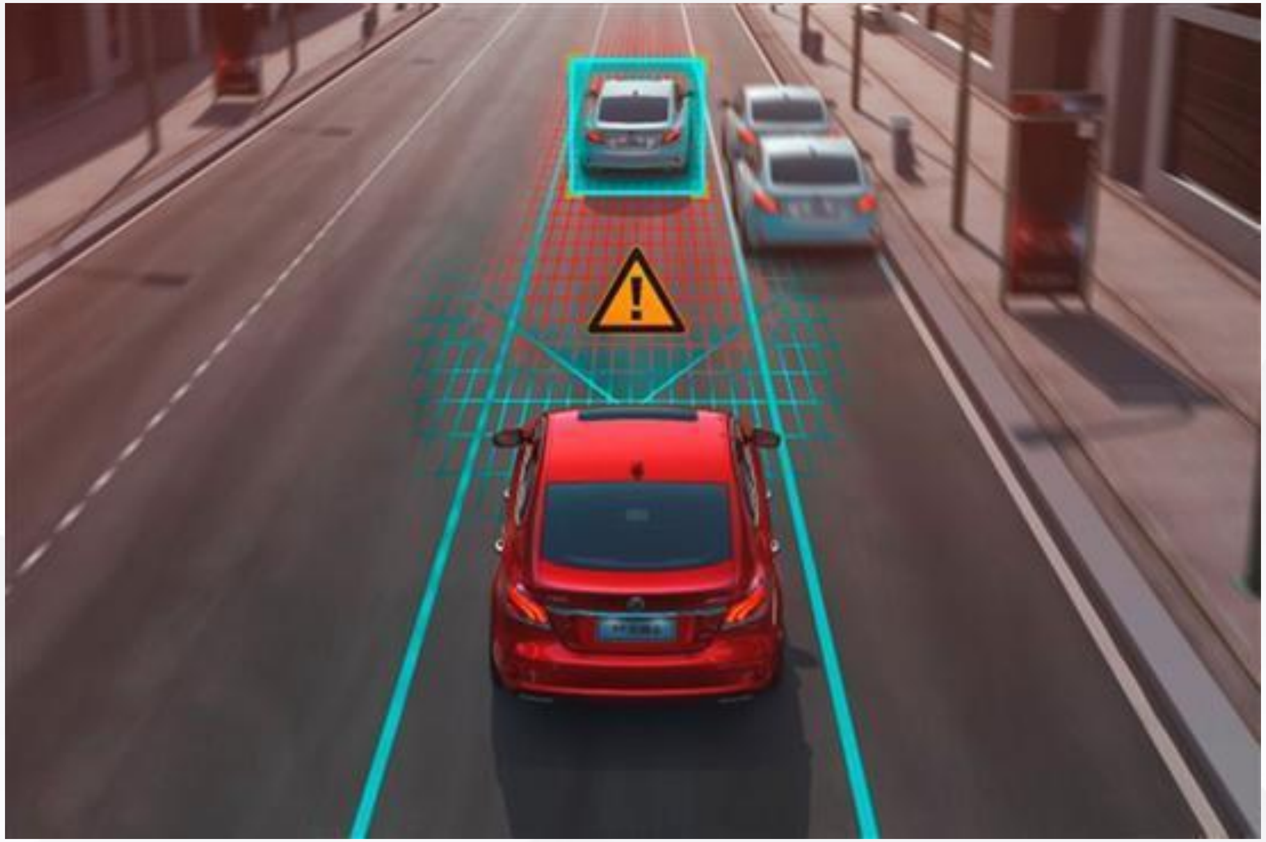
Problems & Contributions

01 Problems

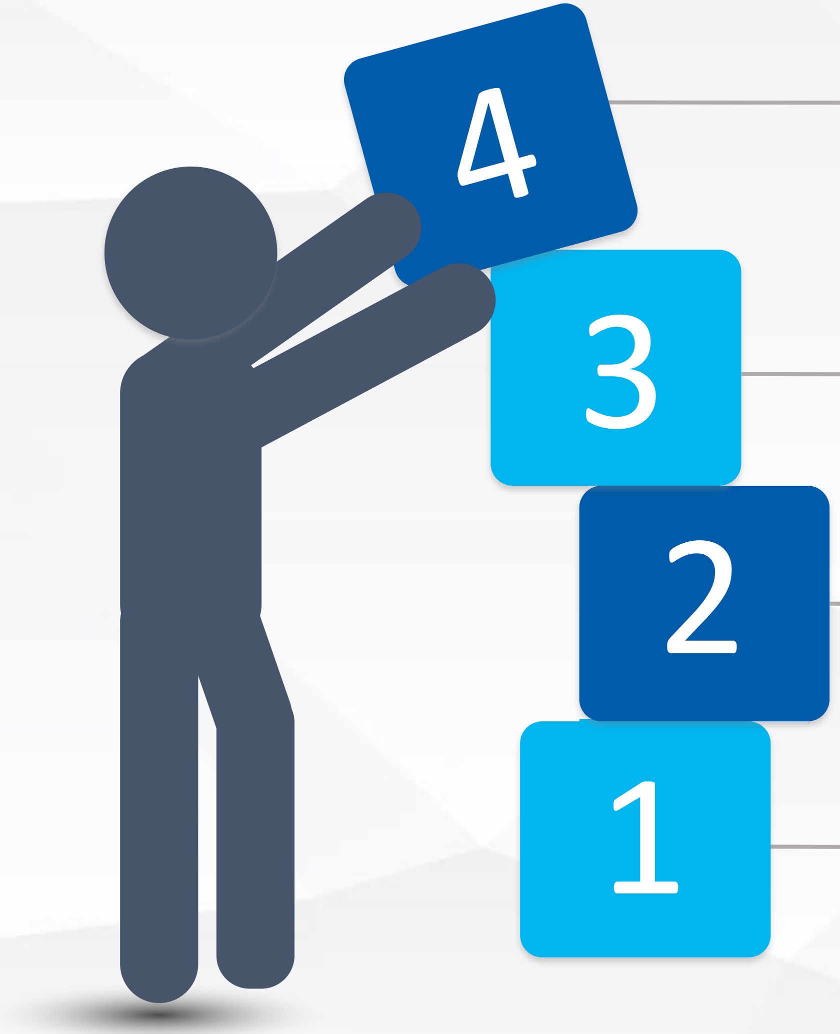
- **Vehicle Collision Warning system (VCW)**
- **quantified by time-to-collision (TTC)**

Two problems:

- **The intensity-based methods cannot handle the challenging situations such as occlusion and rapid scale changes.**
- **Descriptor-based methods can capture the changes, but lack of the accuracy of scale estimation.**



01 Contributions



4

Propose to use an efficient second-order minimization scheme that computes the scale change by minimizing the objective function

3

Introduce a two-stage attention mechanism to handle occlusion and noise corruption.

2

Introduce a method of salient object detection that enables the system to quickly focus on the most relevant target vehicle.

1

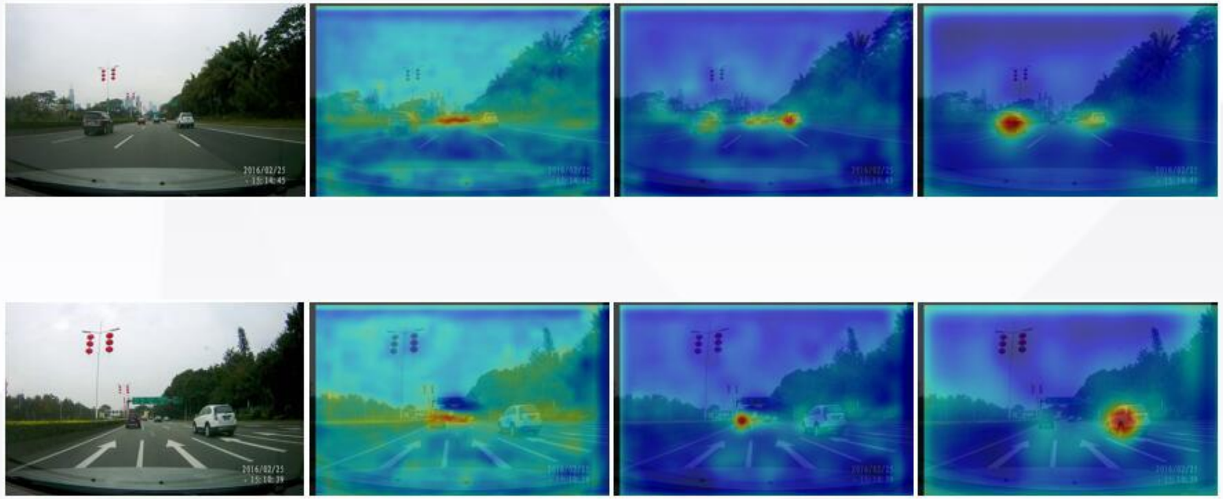
Integrate intensity-based method and descriptor-based method with a regularization term in Lie Algebra.

PART 2

Methods

Part A： Salient Object Detection

The results of salient object detection from the videos captured on the road



From left to right: the original images, the saliency maps of three different depth of layers.

Part B：Descriptor Matching

$$y = \arg\min_{y} \phi_m(y) ,$$

$$\hat{H} = \begin{bmatrix} \hat{s}^{\frac{1}{3}} & 0 & \hat{s}^{-\frac{2}{3}} \cdot \hat{t}_x \\ 0 & \hat{s}^{\frac{1}{3}} & \hat{s}^{-\frac{2}{3}} \cdot \hat{t}_y \\ 0 & 0 & \hat{s}^{-\frac{2}{3}} \end{bmatrix}$$

Part C：Intensity Matching

$$\mathcal{L}_{gray}(x) = \frac{1}{2} \sum_{p \in \Omega} \|\delta(x,p)\|_2^2 ,$$

$$\delta(x,p) = I_t \left(\Psi\left(\hat{H} \cdot H(x), p\right)\right) - I_0(p),$$

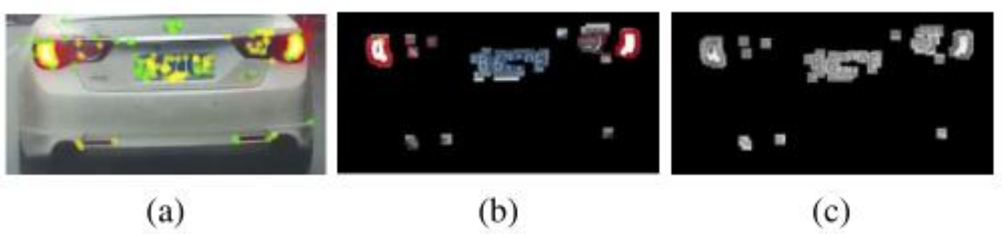
Part D: Intensity Matching with Attention Mechanism

- Two-stage attention mechanism which is composed of hard attention mechanism and soft attention mechanism.
- Hard attention mechanism: A binary mask $M_t(p)$
- Soft attention mechanism: Positive weight $W_t\ (p)$

$$\mathcal{L}_{robust}(x) = \frac{1}{2}\sum_{p\in\Omega} M_t(p) \cdot W_t\ (p)\ \|\delta(x,p)\|_2^2$$

$$\sum_{p\epsilon P} \frac{I_t(\Psi(\widehat{H},p))^T \cdot I_0\ (p)}{\|I_0\ (p)\|_2 \cdot \left\|I_t(\Psi(\widehat{H},p))\ \right\|_2} > \theta_P$$

$$W_t\ (p) = |\delta(x,p)|^{-c}$$



The example of mask and weight matrix function. (a) Key points. (b) Mask. (c) Weight.

Part E&F: Regularization and Optimization

PART 3

Results

03 Results

(a) Bump

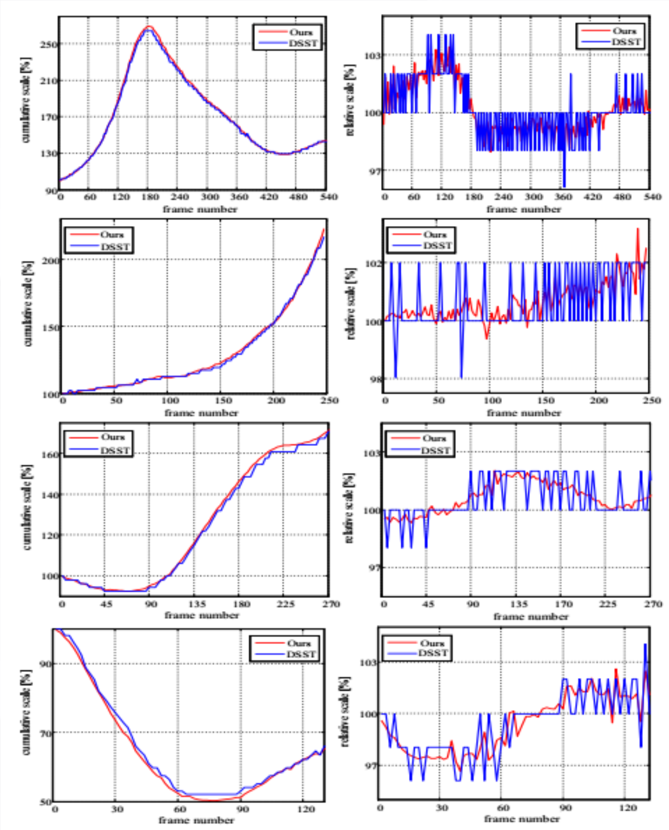
(b) Flashlight

(c) Taillight

(d) Wiper

Tracking results on four representative videos:

- (a) large displacements caused by abrupt bumps of the lead vehicle.
- (b)(c) demonstrate the results at night.
- (d) the target is severely occluded by the wiper and blurred by raindrops.



The relative scale curves and cumulative scale curves of our approach and DSST in 4 real world videos. From top to bottom, Bump, Flashlight, Taillight, Wiper.

PART 4

Funding

Funding

This work is supported by the National Natural Science Foundation of China under Grants 61603022, 61876112 and 61601311, Support Project of High-level Teachers in Beijing Municipal Universities in the Period of 13th Five-year Plan (CIT&TCD20170322).

References

[1] Redmon, Joseph, and Ali Farhadi. Yolov3: An Incremental Improvement. In CVPR ’ 2018
[2] Danelljan, Martin and Häger, Gustav and Khan, Fahad and Felsberg, Michael. Accurate Scale Estimation for Robust Visual Tracking. In BMVC ’2014
[3] S. Benhimane and E. Malis, “Real-time image-based tracking of planes using efficient second-order minimization,” in IEEE International Conference on Intelligent Robots and Systems, vol. 1, 2004, pp. 943–948.

Capital Normal University

Thanks For Listening!