Improving the Accuracy of Stereo Visual Odometry Using Visual Illumination Estimation

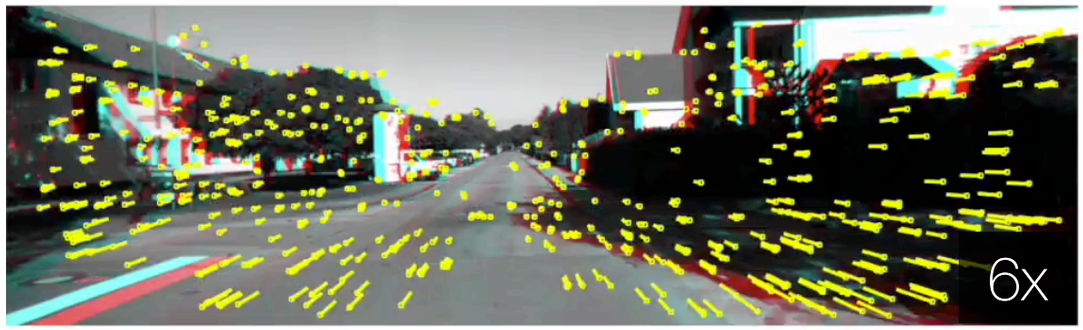
Lee Clement, Valentin Peretroukhin, and Jonathan Kelly

International Symposium on Experimental Robotics (ISER 2016)
Tokyo, Japan

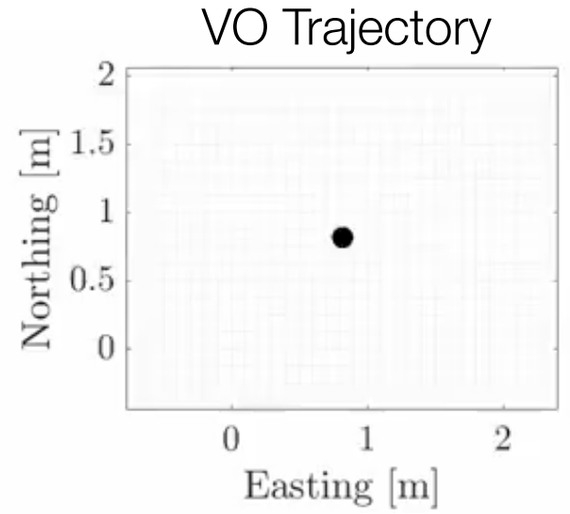
4 October 2016

Stereo Visual Odometry (VO)

Goal: Estimate the egomotion of a moving platform without GPS

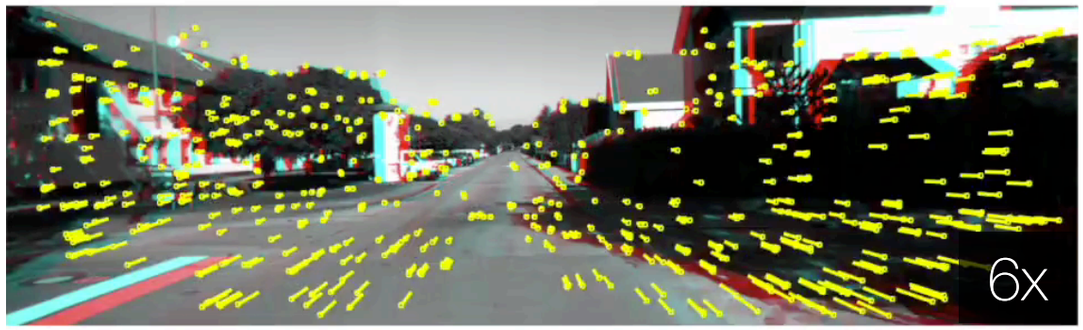


A. Geiger et al., “Vision meets robotics: The KITTI dataset,” IJRR 2013.

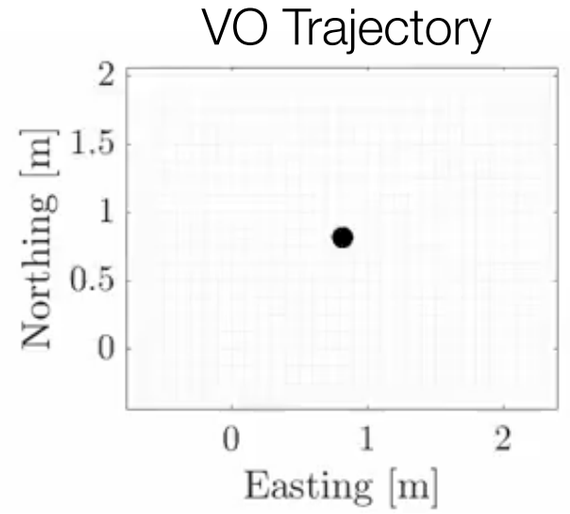


Stereo Visual Odometry (VO)

Goal: Estimate the egomotion of a moving platform without GPS

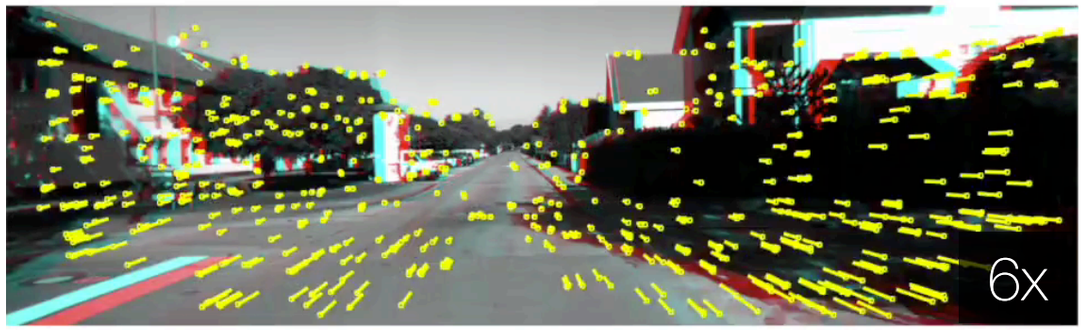


A. Geiger et al., “Vision meets robotics: The KITTI dataset,” IJRR 2013.

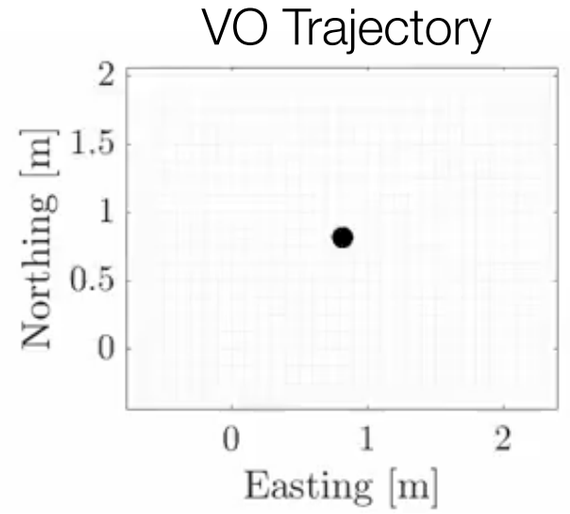


Stereo Visual Odometry (VO)

Goal: Estimate the egomotion of a moving platform without GPS



A. Geiger et al., “Vision meets robotics: The KITTI dataset,” IJRR 2013.



Visual Odometry	vs.	Wheel Odometry
✔ Immune to wheel slip		✘ Fails in slippery terrain
✔ Any type of vehicle		✘ Ground vehicles only
✔ Loop closures		✘ No loop closures

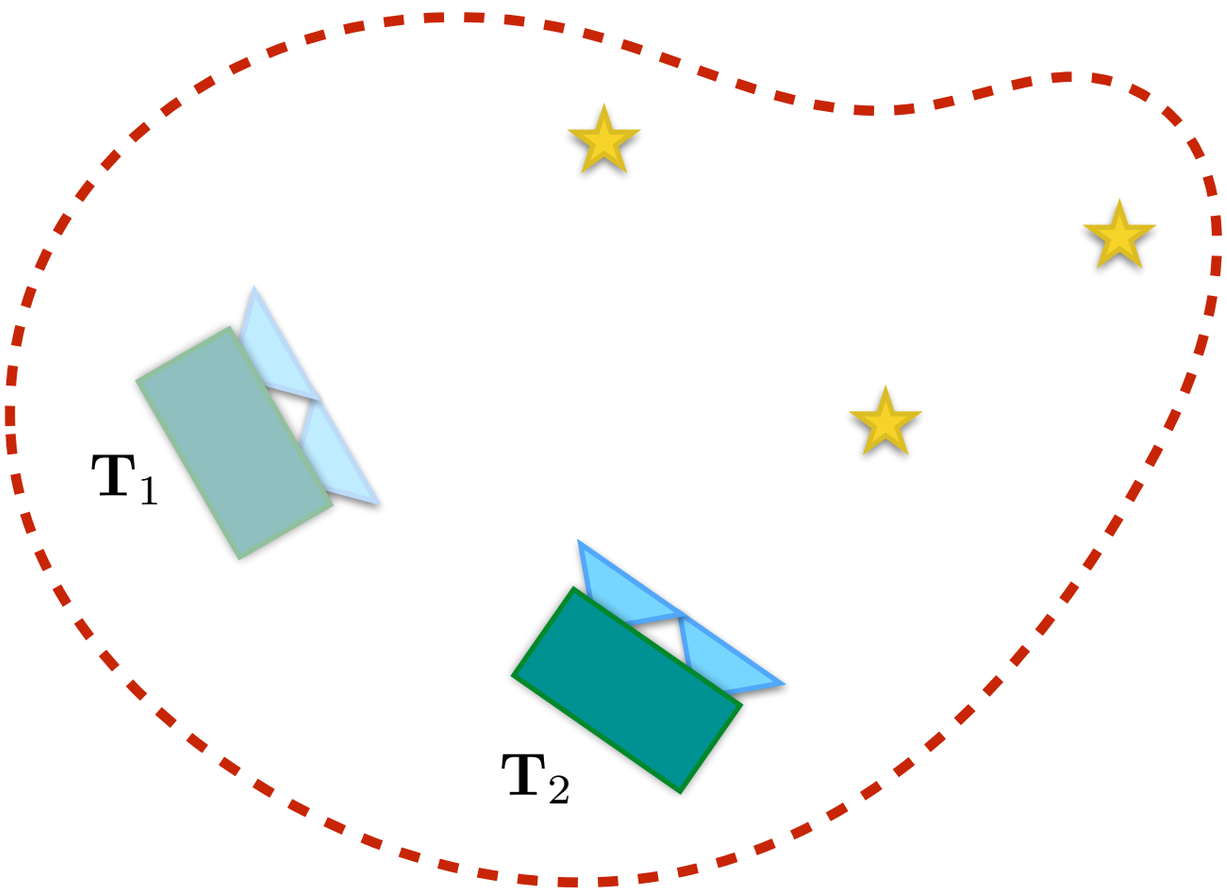
Sliding Window Stereo VO

Estimate a **sliding window** of SE(3) poses by **tracking keypoints**

Sliding Window Stereo VO

Estimate a **sliding window** of SE(3) poses by **tracking keypoints**

Formulate as an optimization problem to **minimize reprojection error**



$$\text{Cost} = \sum_{k} \sum_{j} \mathbf{e}_{\mathbf{y}_{k,j}}^{T} \mathbf{R}_{\mathbf{y}_{k,j}}^{-1} \mathbf{e}_{\mathbf{y}_{k,j}}$$

(to minimize)

Reprojection error

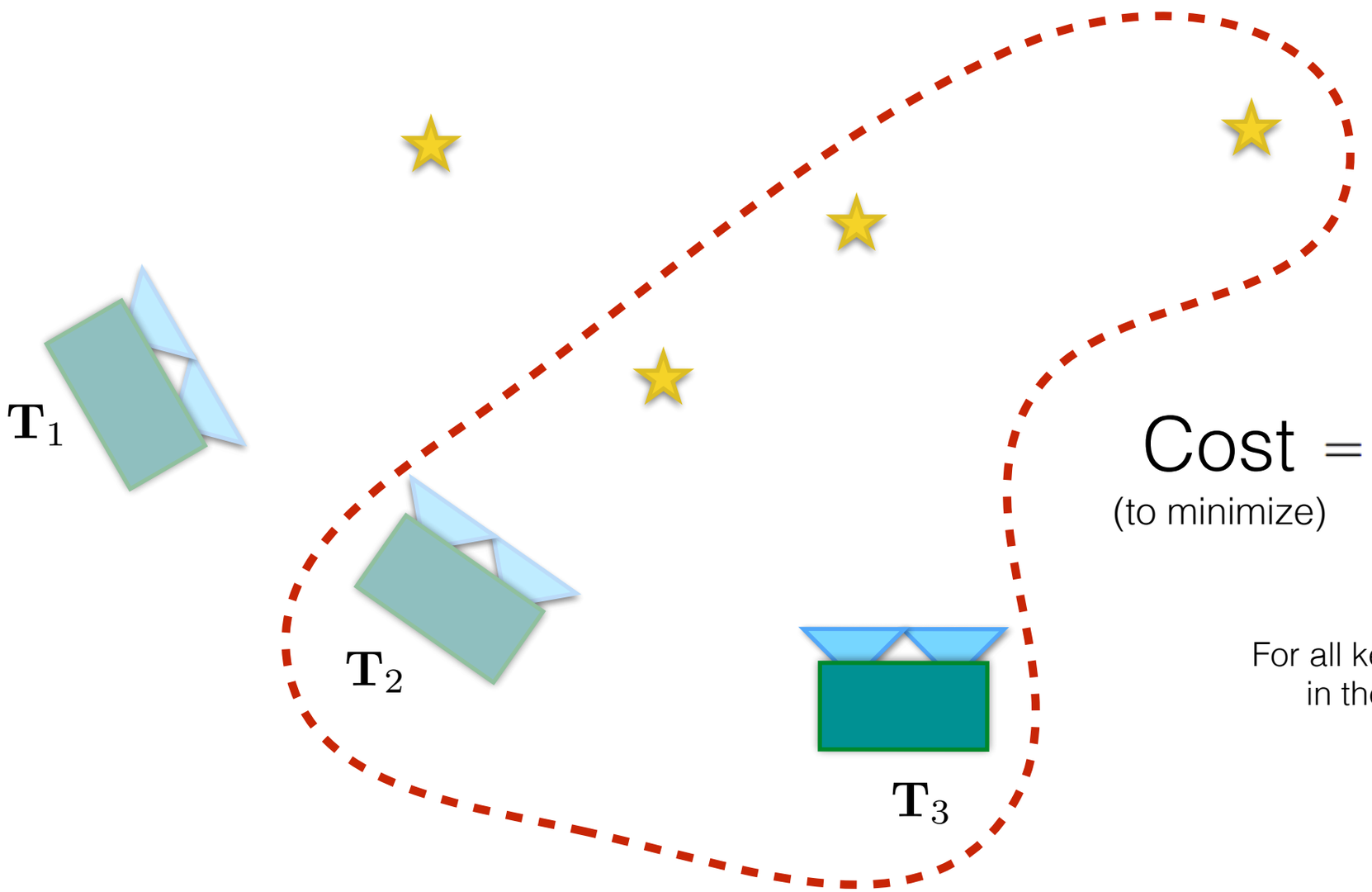
Weight

For all keypoints and poses in the active window

Sliding Window Stereo VO

Estimate a **sliding window** of SE(3) poses by **tracking keypoints**

Formulate as an optimization problem to **minimize reprojection error**



Reprojection error

$$\text{Cost} = \sum_{k} \sum_{j} \mathbf{e}_{\mathbf{y}_{k,j}}^{T} \mathbf{R}_{\mathbf{y}_{k,j}}^{-1} \mathbf{e}_{\mathbf{y}_{k,j}}$$

(to minimize)

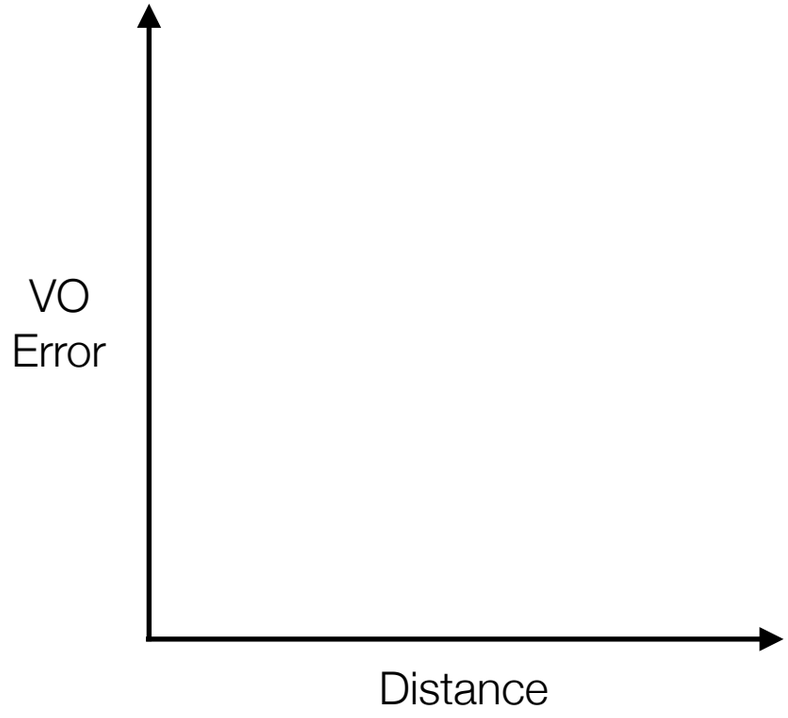
Weight

For all keypoints and poses in the active window

Sun-Aided Stereo VO

VO is a dead-reckoning technique and suffers from **superlinear error growth**, largely due to **accumulated orientation error**

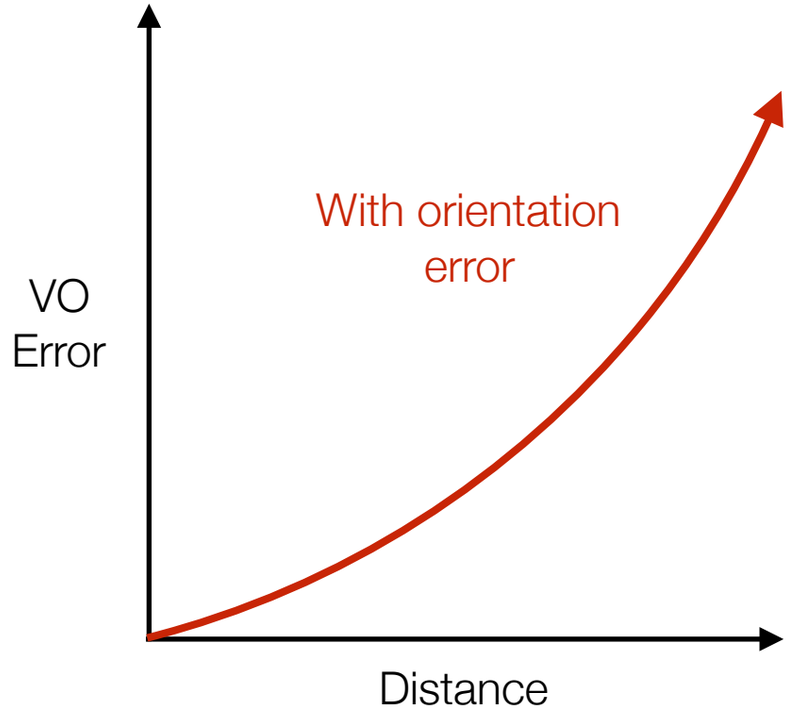
C. F. Olson et al. “Rover navigation using stereo ego-motion,” Rob. Auton. Syst., 2003



Sun-Aided Stereo VO

VO is a dead-reckoning technique and suffers from **superlinear error growth**, largely due to **accumulated orientation error**

C. F. Olson et al. “Rover navigation using stereo ego-motion,” Rob. Auton. Syst., 2003

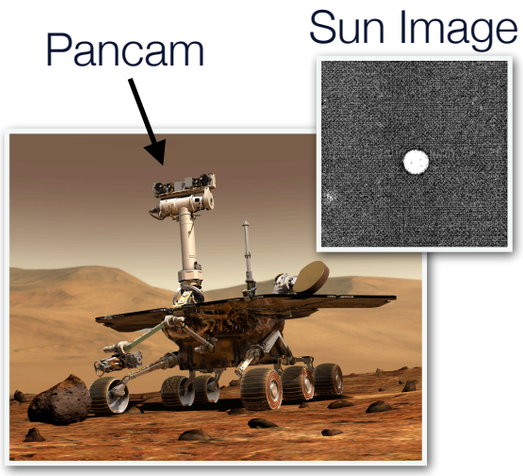


Sun-Aided Stereo VO

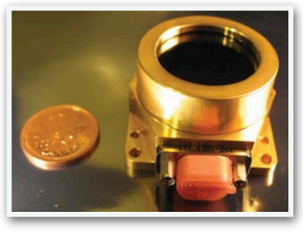
VO is a dead-reckoning technique and suffers from **superlinear error growth**, largely due to **accumulated orientation error**

C. F. Olson et al. “Rover navigation using stereo ego-motion,” Rob. Auton. Syst., 2003

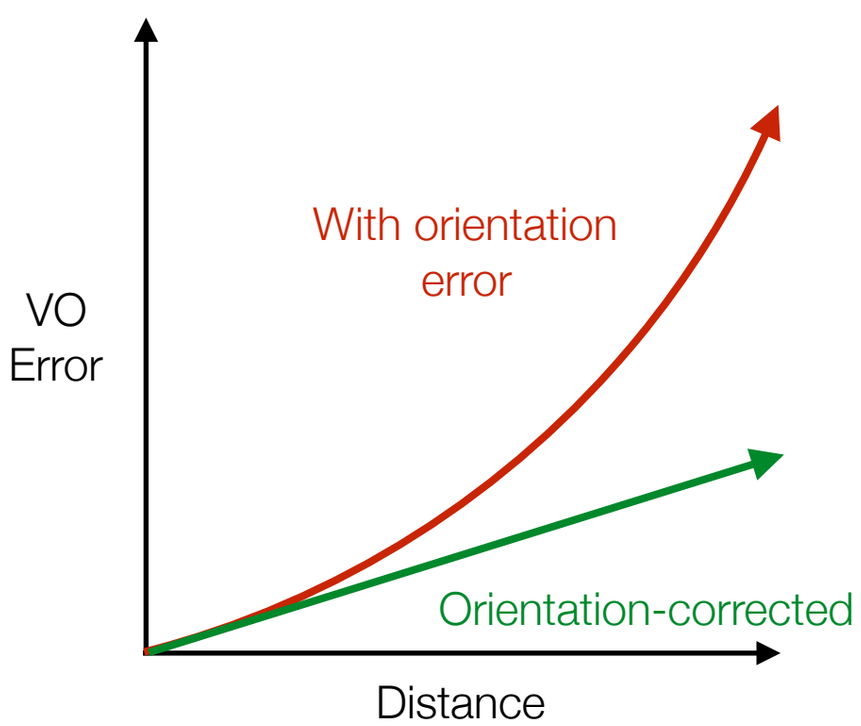
Drift can be reduced using **absolute orientation information** from a sun sensor



Source: NASA; A. R. Eisenman et al., “Sun sensing on the Mars exploration rovers,” in Aerosp. Conf. Proc. 2002



Sinclair Interplanetary SS-411 digital sun sensor

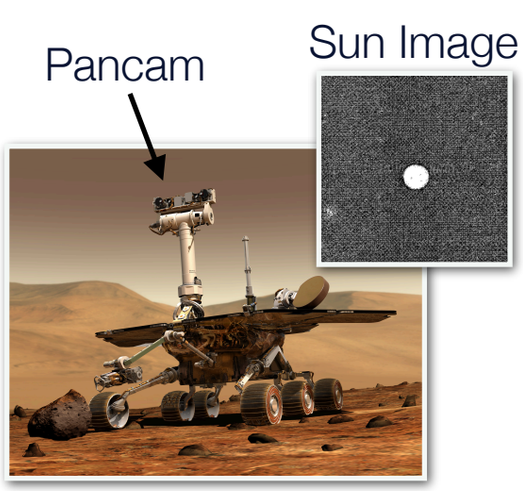


Sun-Aided Stereo VO

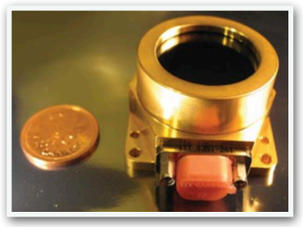
VO is a dead-reckoning technique and suffers from **superlinear error growth**, largely due to **accumulated orientation error**

C. F. Olson et al. “Rover navigation using stereo ego-motion,” Rob. Auton. Syst., 2003

Drift can be reduced using **absolute orientation information** from a sun sensor



Source: NASA; A. R. Eisenman et al., “Sun sensing on the Mars exploration rovers,” in Aerosp. Conf. Proc. 2002



Sinclair Interplanetary SS-411 digital sun sensor

$$\text{Cost} = \sum_k \left(\sum_j \mathbf{e}_{\mathbf{y}_{k,j}}^T \mathbf{R}_{\mathbf{y}_{k,j}}^{-1} \mathbf{e}_{\mathbf{y}_{k,j}} + \mathbf{e}_{\mathbf{s}_k}^T \mathbf{R}_{\mathbf{s}_k}^{-1} \mathbf{e}_{\mathbf{s}_k} \right)$$

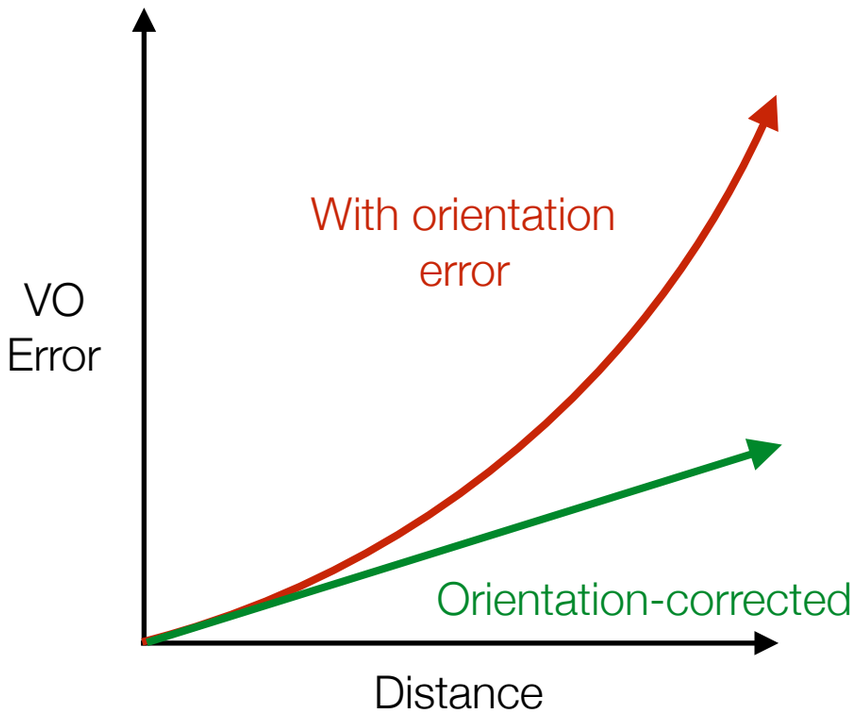
(to minimize)

Reprojection error

Reprojection error weight

Sun error weight

Sun sensor error



A. Lambert, et al., “Field testing of visual odometry aided by a sun sensor and inclinometer,” JFR 2012

Simultaneous Localization and… Sun Sensing?

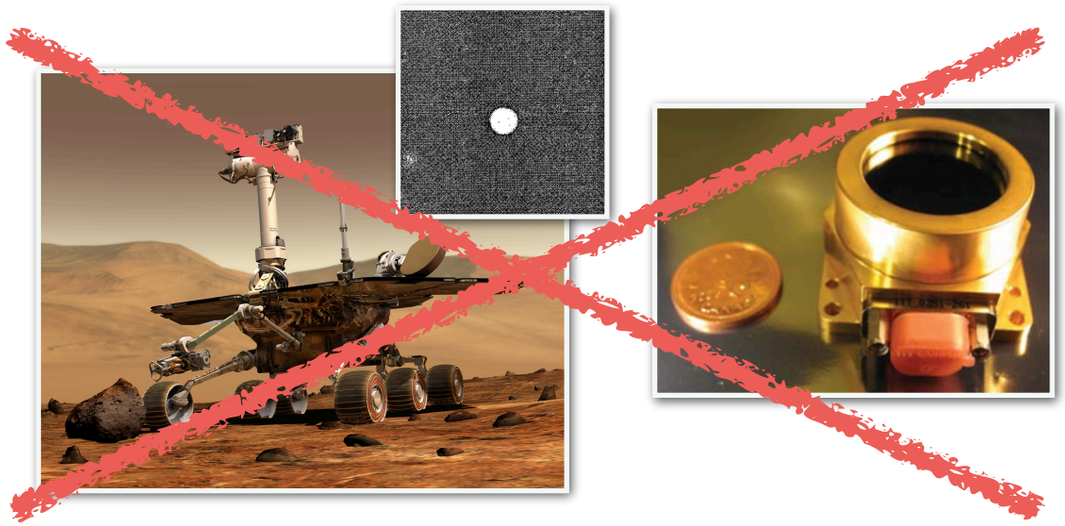
Do we really need a **hardware sun sensor** or **specially oriented camera?**

In other words, do we need to look at the sun to see the sun?

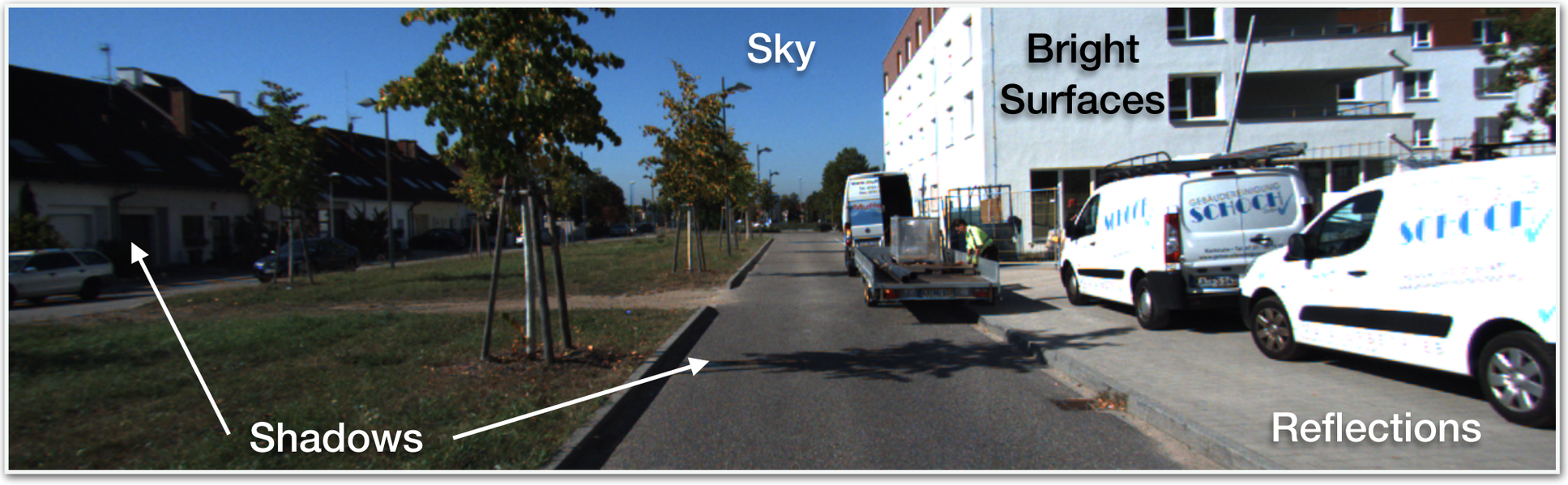
Simultaneous Localization and… Sun Sensing?

Do we really need a **hardware sun sensor** or **specially oriented camera?**

In other words, do we need to look at the sun to see the sun?



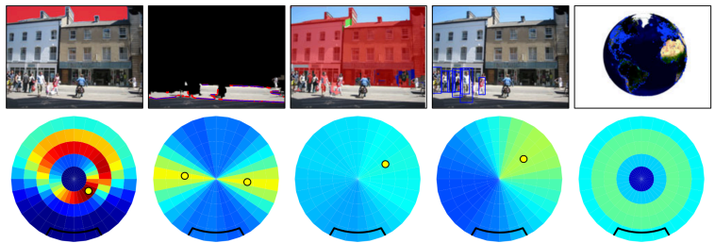
We can **infer** the likely direction of the sun by looking at **visual cues**



Inferring the Sun Direction from Visual Cues

We compare…

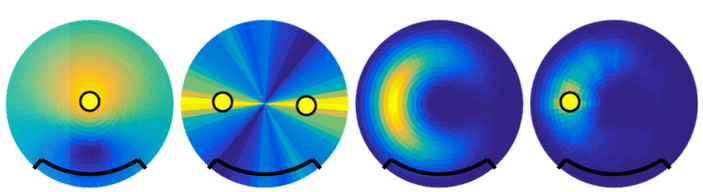
Lalonde



Handcrafted visual cues

J.-F. Lalonde et al., “Estimating the Natural Illumination Conditions from a Single Outdoor Image,” IJCV 2011.

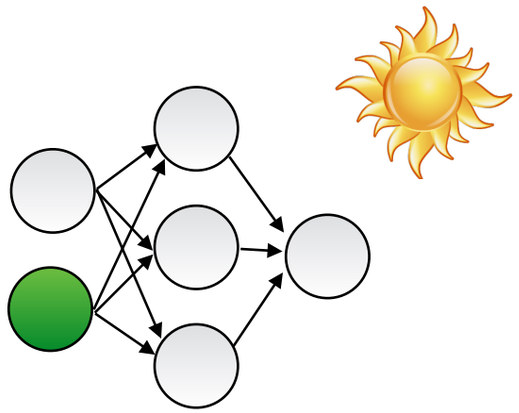
Lalonde-VO



Handcrafted visual cues
+
Information from VO

New in this paper

Sun-CNN



Learn from data

W.-C. Ma et al., “Find your Way by Observing the Sun and Other Semantic Cues,” arXiv [cs.CV], 23-Jun-2016.

Lalonde: Sky, shadows, and probability

Input Image

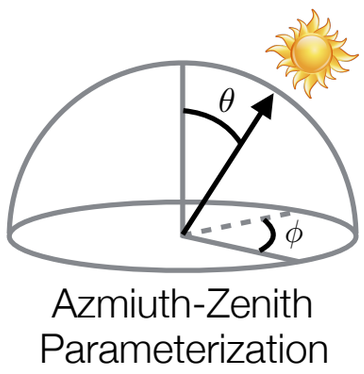
J.-F. Lalonde et al., “Estimating the Natural Illumination Conditions from a Single Outdoor Image,” IJCV 2011.

Lalonde: Sky, shadows, and probability

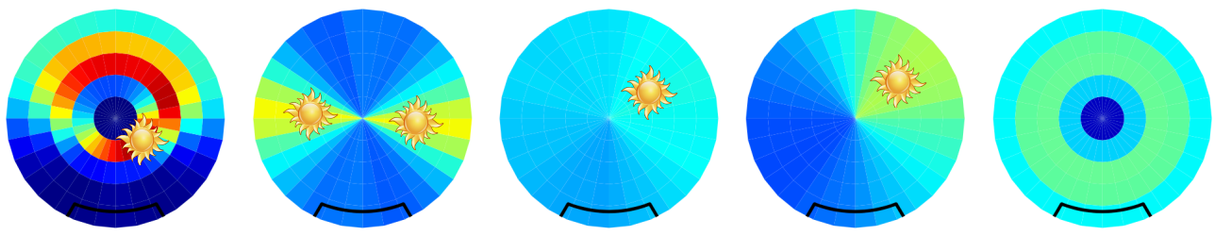
Input Image

Visual Cue Detection

Per-cue Sun Probability



Azmiuth-Zenith Parameterization



Sky model | Shadows | Surfaces | Pedestrians | Prior

Low High

Maximum Likelihood

J.-F. Lalonde et al., “Estimating the Natural Illumination Conditions from a Single Outdoor Image,” IJCV 2011.

Lalonde: Sky, shadows, and probability

Input Image

Visual Cue Detection

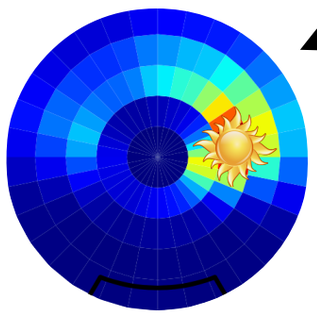
Azmiuth-Zenith Parameterization

Sky model | Shadows | Surfaces | Pedestrians | Prior

Per-cue Sun Probability

Low High

Maximum Likelihood



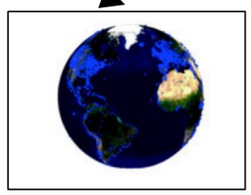
Combined Result

J.-F. Lalonde et al., “Estimating the Natural Illumination Conditions from a Single Outdoor Image,” IJCV 2011.

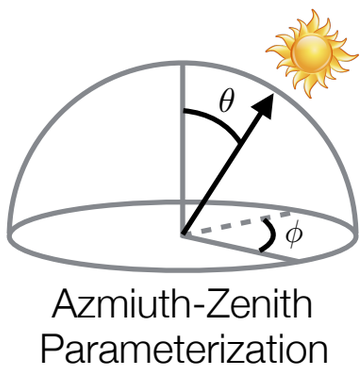
Lalonde: Sky, shadows, and probability

Input Image

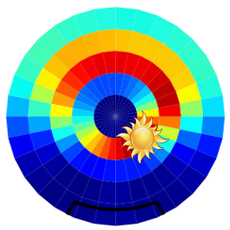
Visual Cue Detection



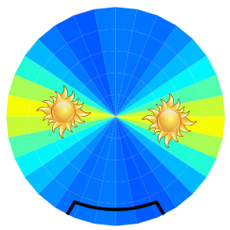
Per-cue Sun Probability



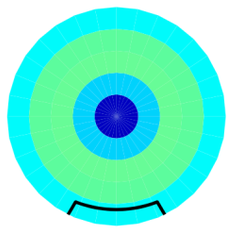
Azmiuth-Zenith Parameterization



Sky model



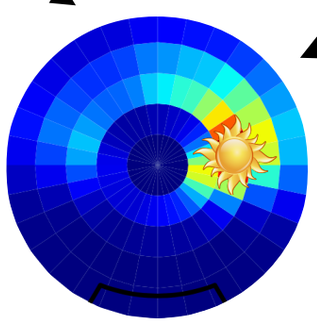
Shadows



Prior

Low High

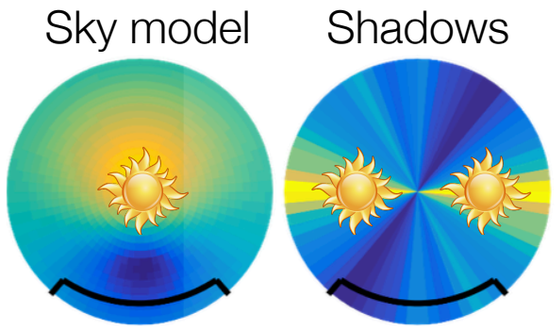
Maximum Likelihood

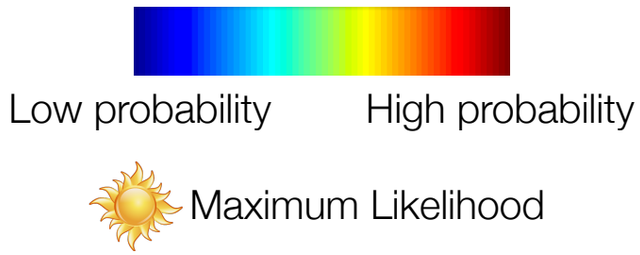


Combined Result

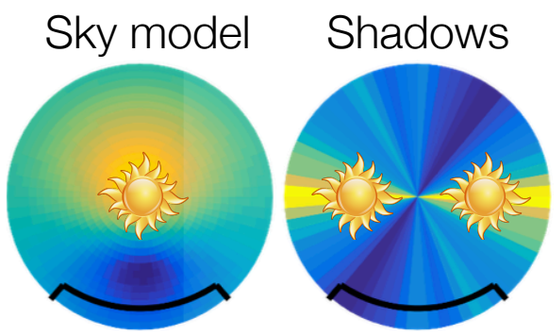
J.-F. Lalonde et al., “Estimating the Natural Illumination Conditions from a Single Outdoor Image,” IJCV 2011.

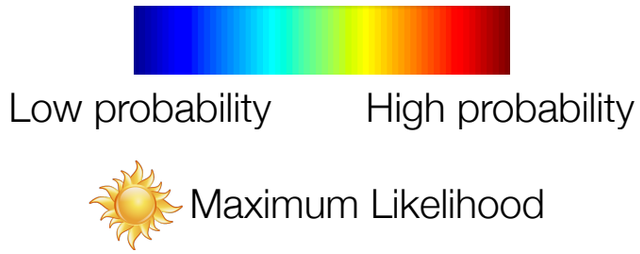
Lalonde-VO: An odometry-informed prior



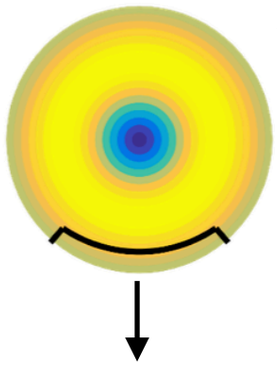


Lalonde-VO: An odometry-informed prior



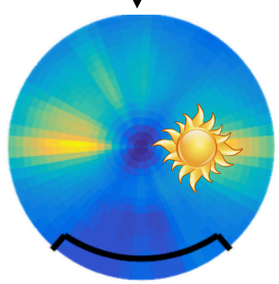


Uniform Prior
(Lalonde)

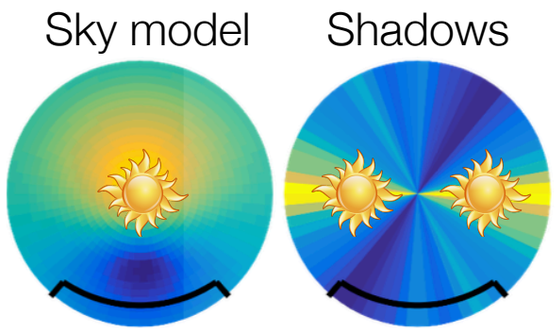


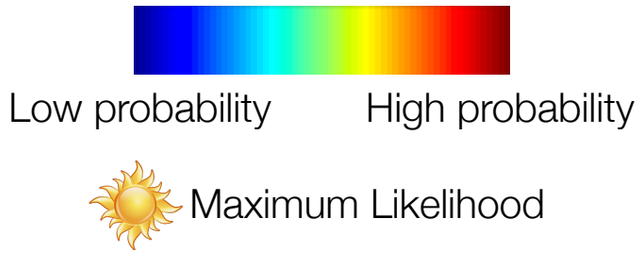
Ambiguous cues generate
outlier observations

Combined
(Incorrect)

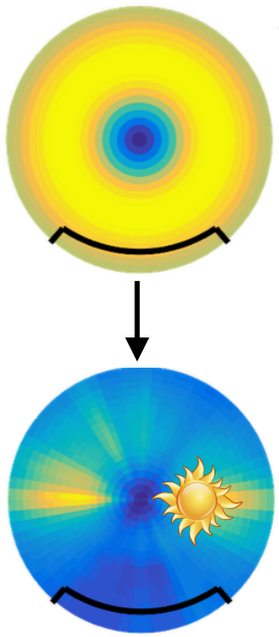


Lalonde-VO: An odometry-informed prior





Uniform Prior
(Lalonde)

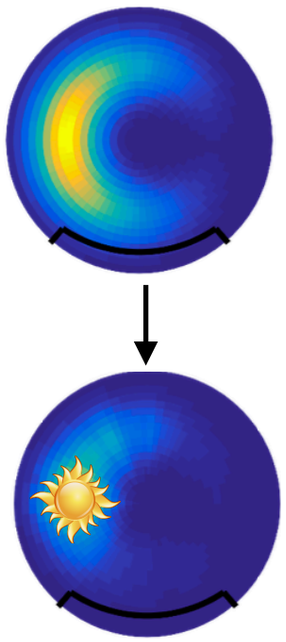


Combined
(Incorrect)

Ambiguous cues generate
outlier observations

Resolve ambiguities
using information from VO!

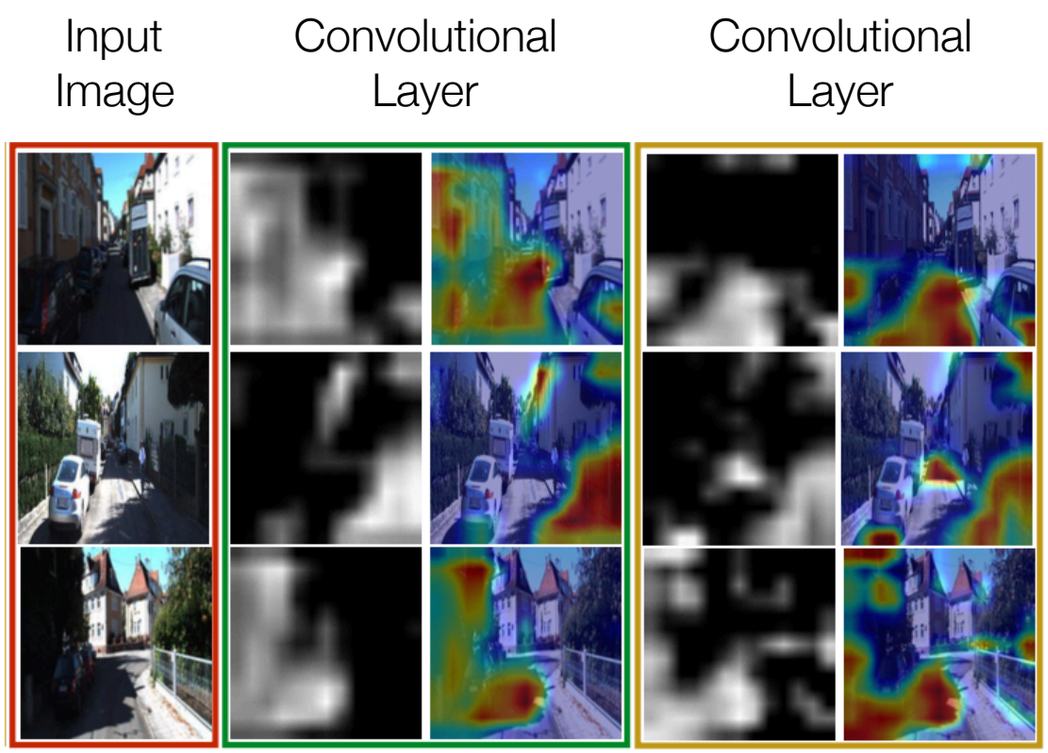
VO-informed Prior
(Lalonde-VO)



Combined
(Correct)

Sun-CNN: Learning to find the sun

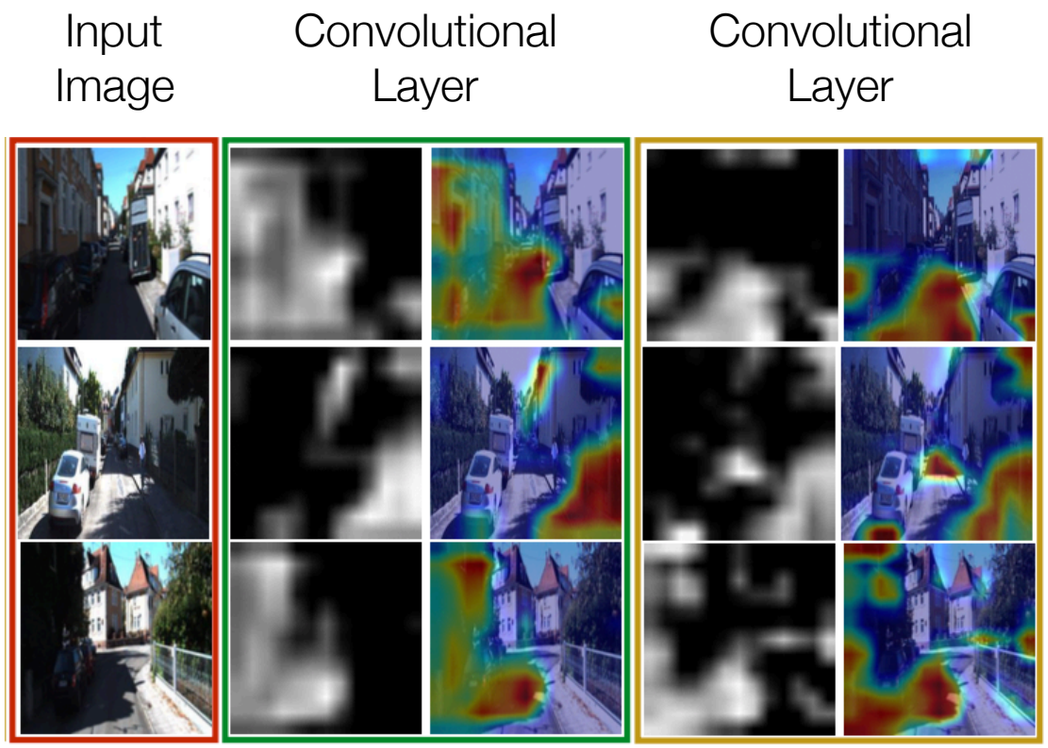
Handcrafted feature detectors often do not generalize well.
Why not **learn** cues from data using a **deep neural network?**



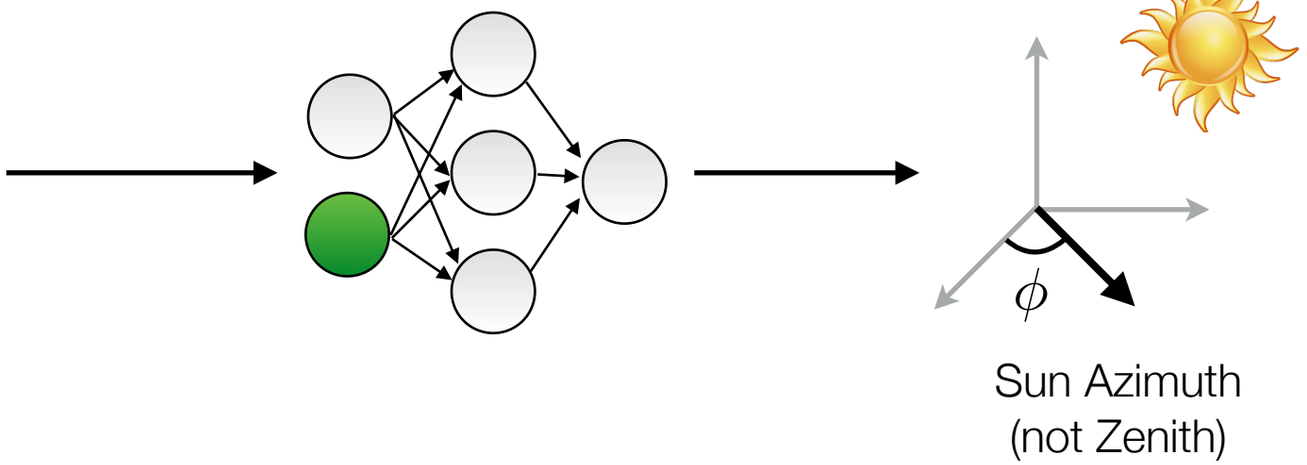
W.-C. Ma et al., “Find your Way by Observing the Sun and Other Semantic Cues,” arXiv [cs.CV], 23-Jun-2016.

Sun-CNN: Learning to find the sun

Handcrafted feature detectors often do not generalize well.
Why not **learn** cues from data using a **deep neural network?**



AlexNet pre-trained on ImageNet
Additional training on KITTI odometry dataset



W.-C. Ma et al., "Find your Way by Observing the Sun and Other Semantic Cues," arXiv [cs.CV], 23-Jun-2016.

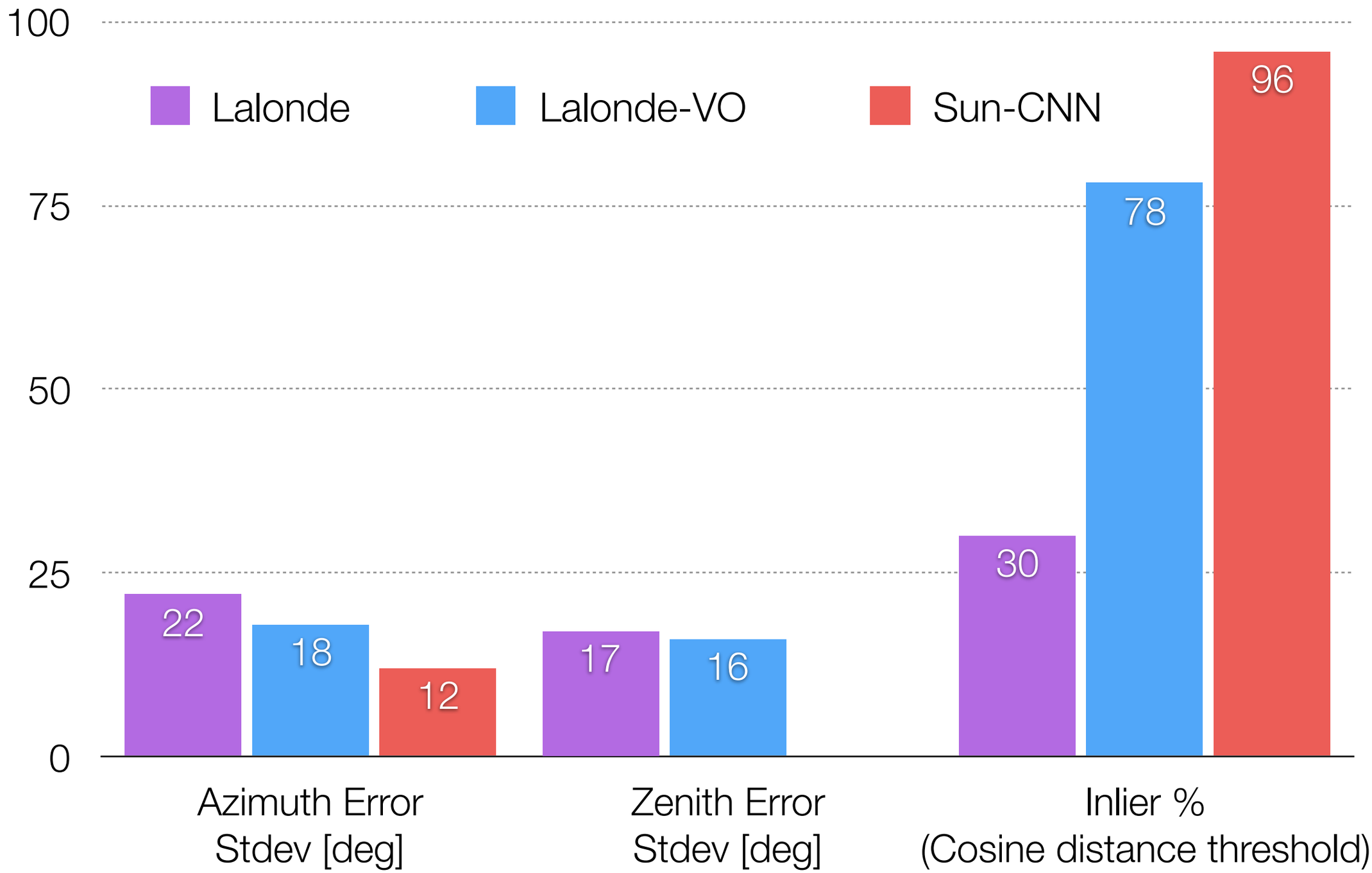
Experiments: The KITTI Dataset

We compare stereo VO accuracy on a combined **7.8 km of urban driving** using

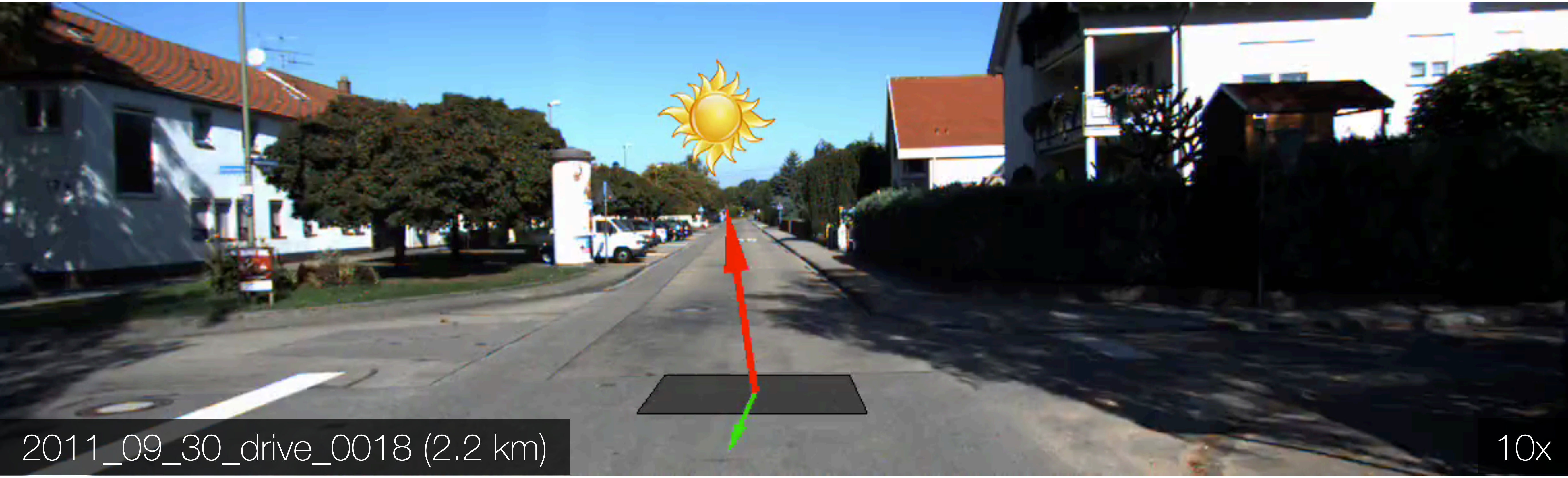
- No sun detection
- Lalonde
- Lalonde-VO
- Sun-CNN

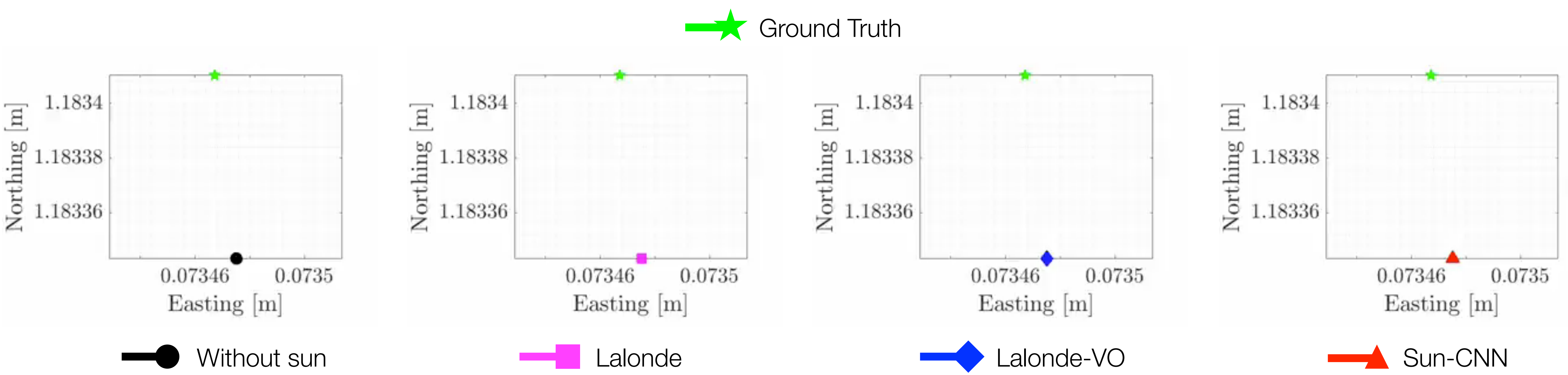
A. Geiger et al., “Vision meets robotics: The KITTI dataset,” IJRR 2013.

Experiments: The KITTI Dataset

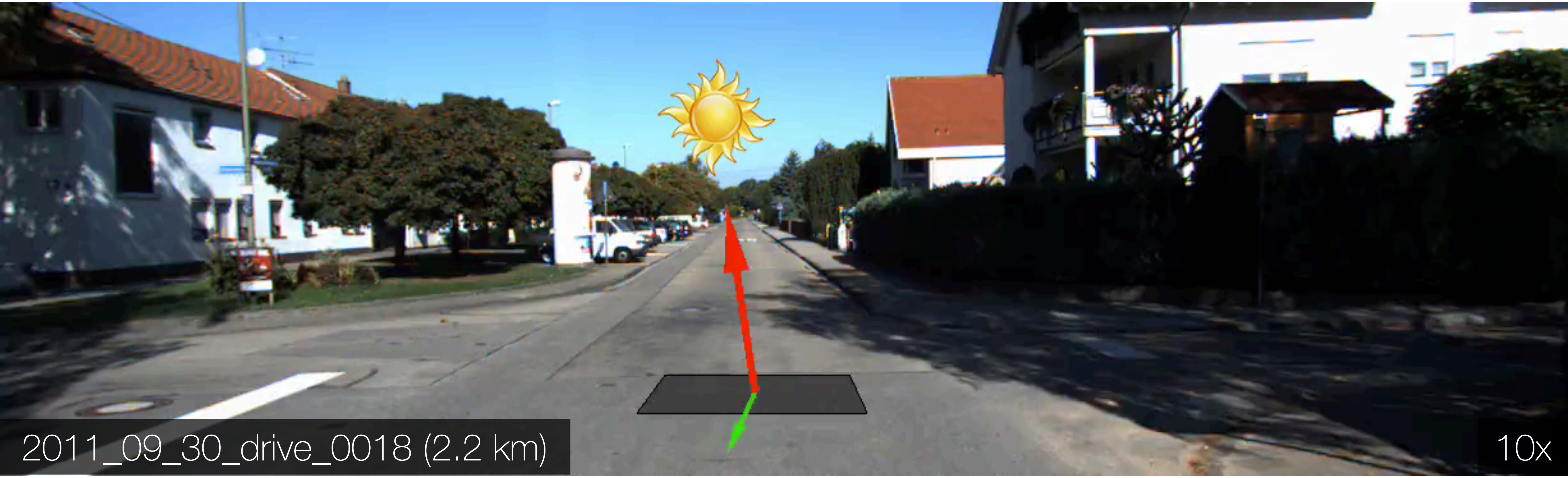


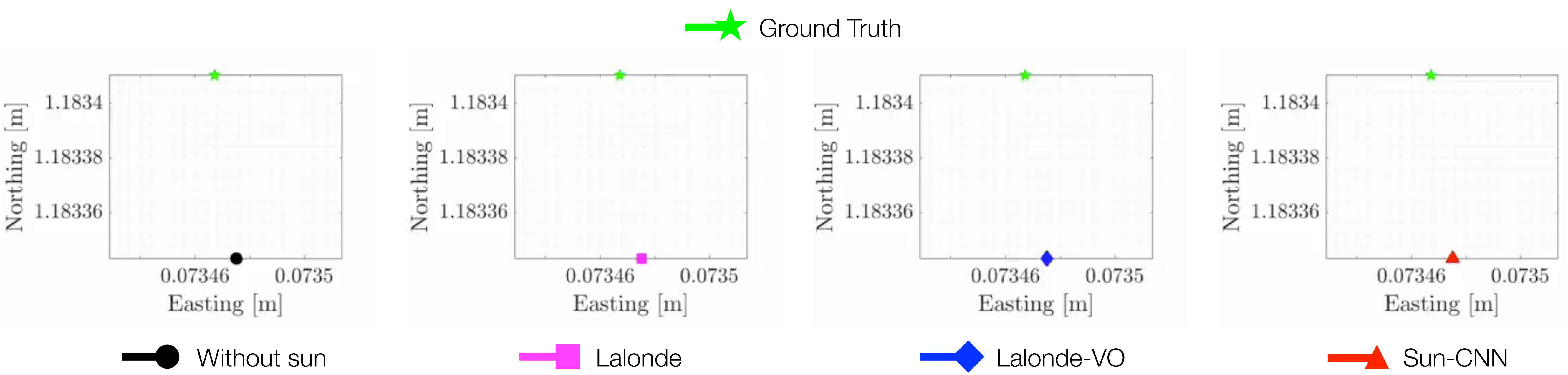
Experiments: The KITTI Dataset



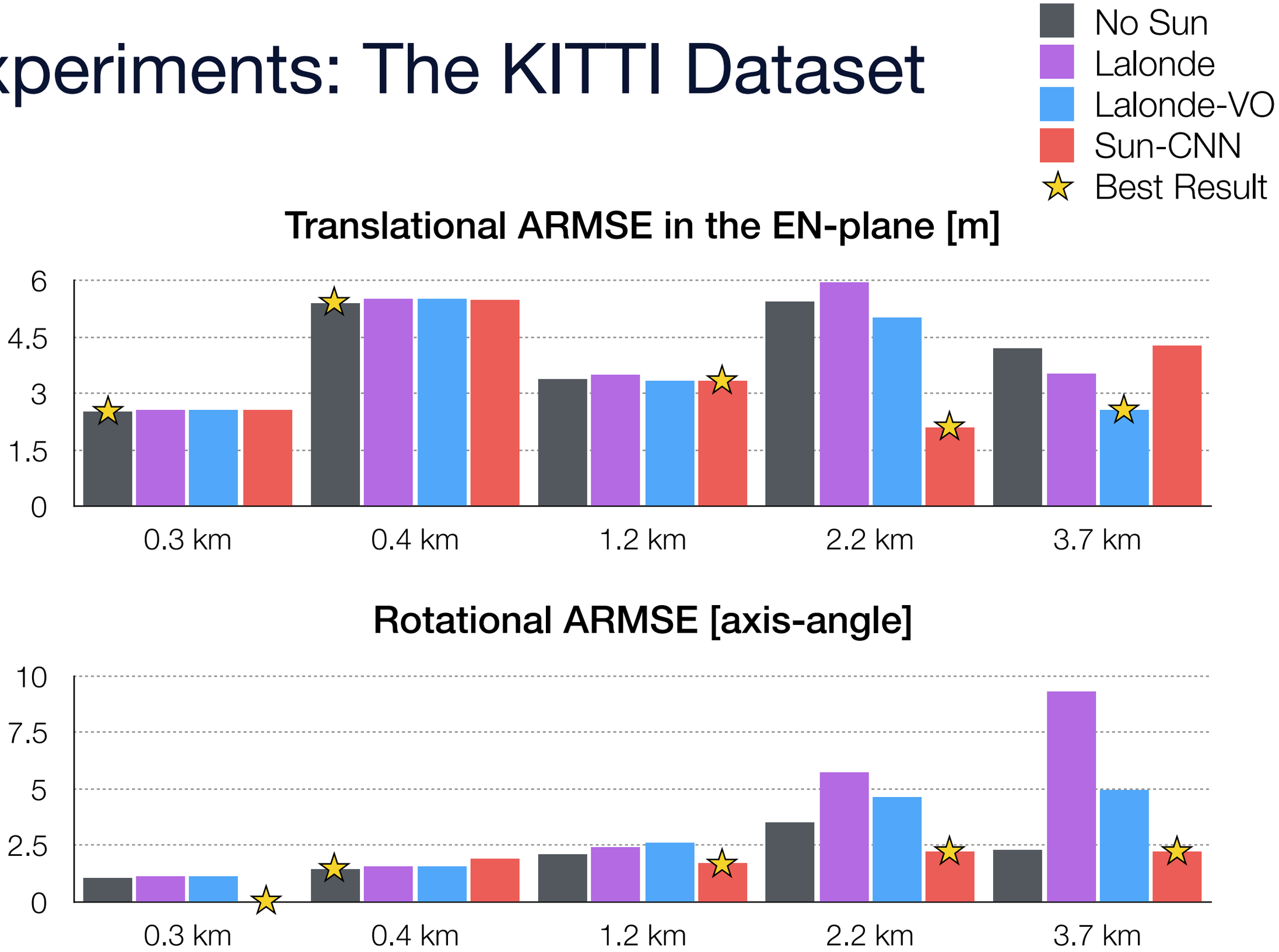


Experiments: The KITTI Dataset





Experiments: The KITTI Dataset

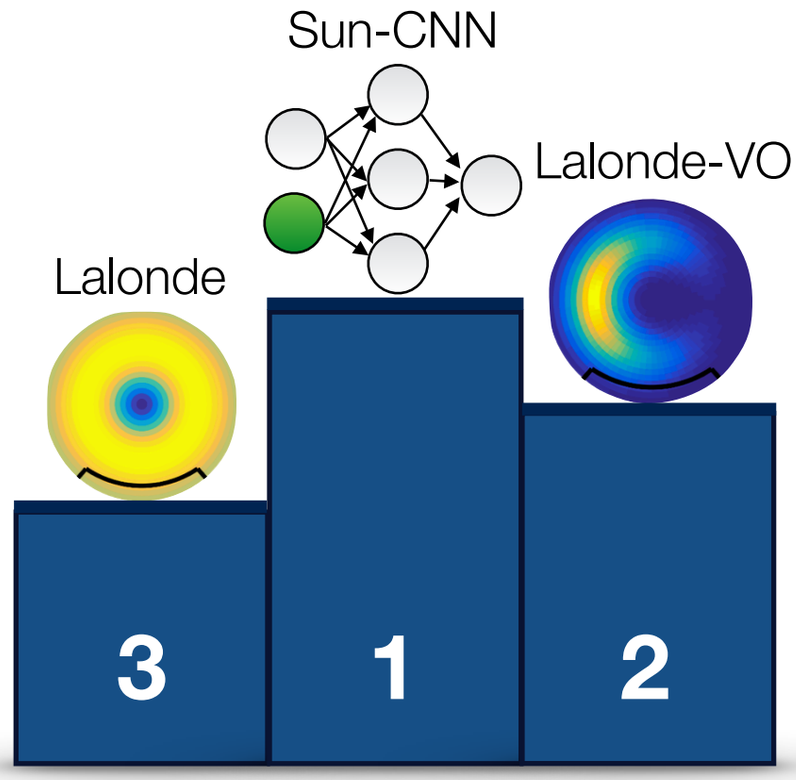


Conclusions

We can **reduce estimation error** in stereo visual odometry by exploiting **global illumination information** already available in the existing image stream — **no additional hardware** required

Hand-crafted visual cues are less reliable than cues learned from data using a Convolutional Neural Network

We recommend **Sun-CNN** for this application for now, but…



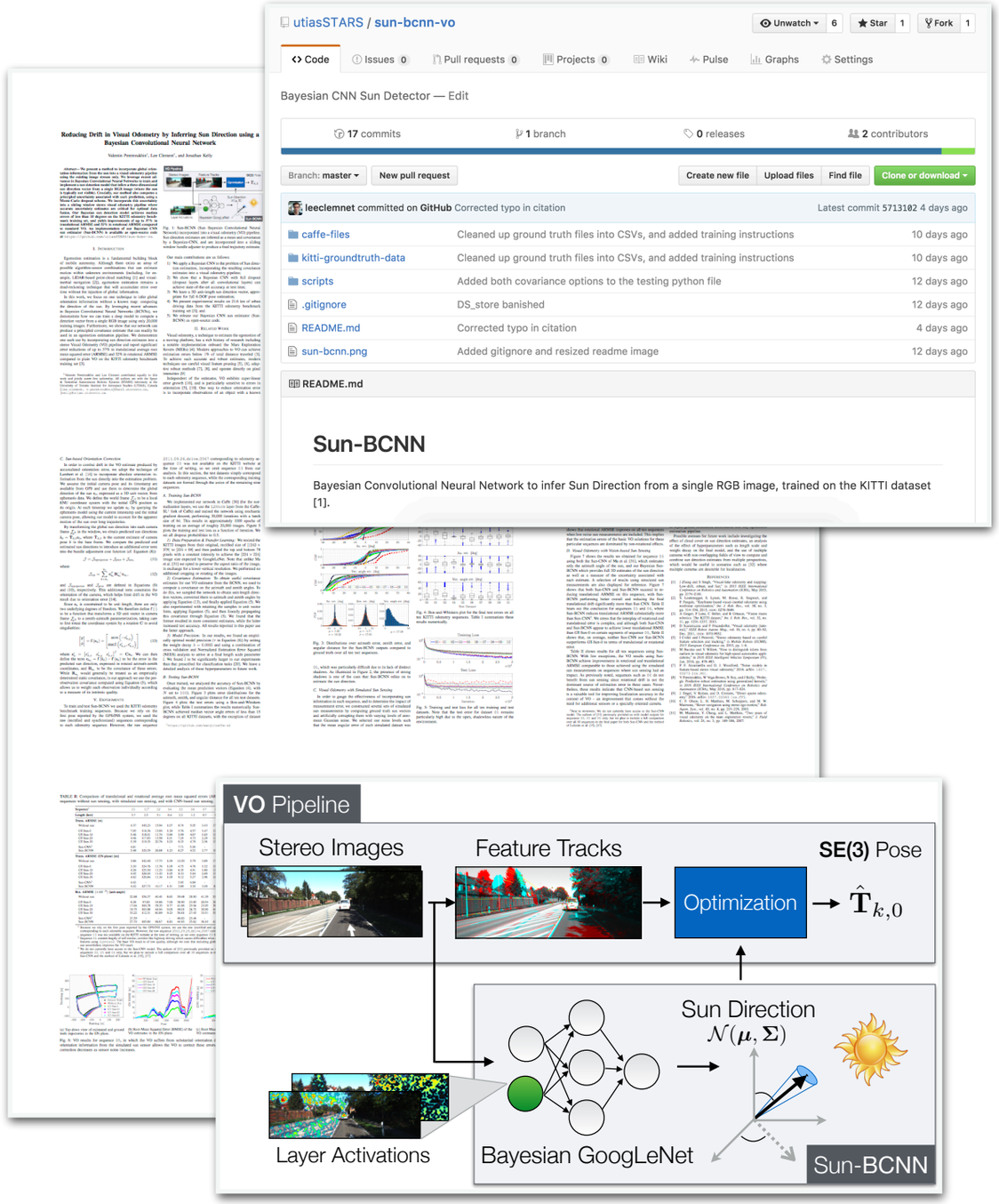
New Work: Sun Estimation with Bayesian CNNs

Sun-BCNN

- Bayesian CNN — principled covariances for all measurements
- Alternative network architecture and transfer learning (based on PoseNet)
  A. Kendall and R. Cipolla, "Modelling uncertainty in deep learning for camera relocalization," ICRA 2016.
- Full 3D sun vector estimation

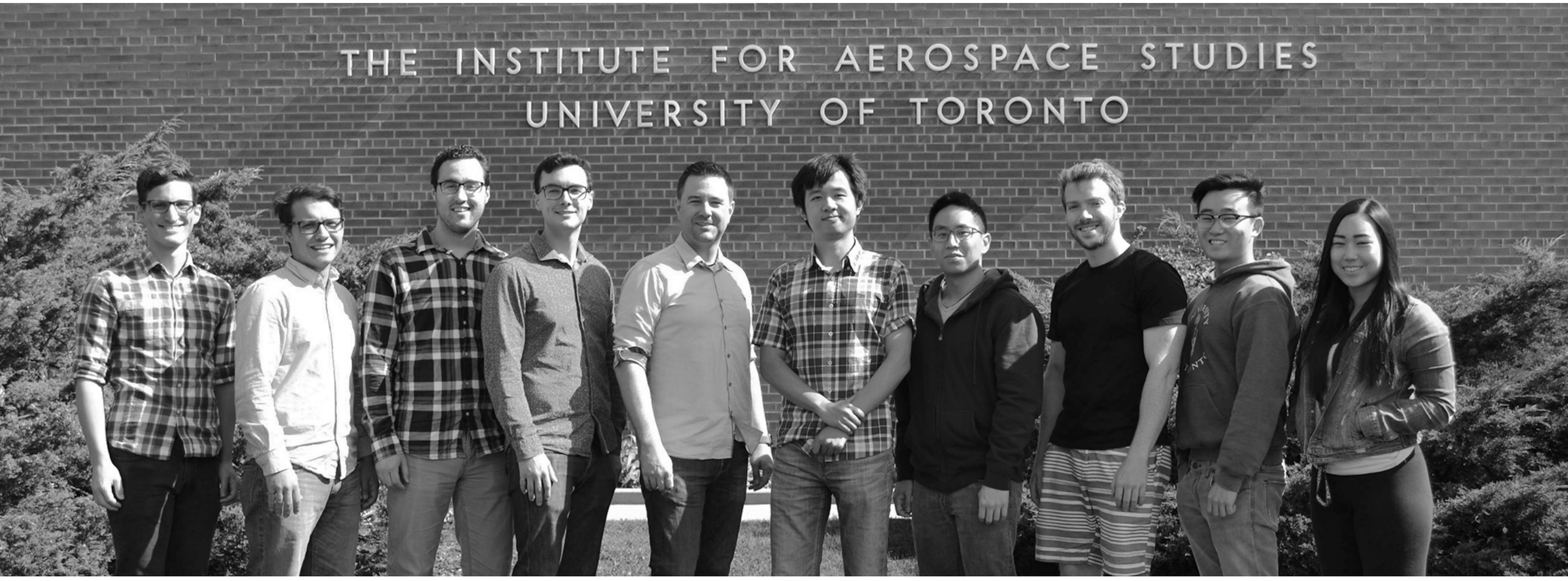
ICRA Submission: arxiv.org/abs/1609.05993

Code: github.com/utiasSTARS/sun-bcnn-vo



Thank you!

Questions?



Lee (Me!) **Valentin** **Jonathan**

UNIVERSITY OF TORONTO

lee.clement@mail.utoronto.ca

http://starslab.ca

STARS LABORATORY