Optimizing Camera Perspective for Stereo Visual Odometry

Valentin Peretroukhin[†]
Jonathan Kelly[†]
Timothy D. Barfoot[*]

University of Toronto
Institute for Aerospace Studies

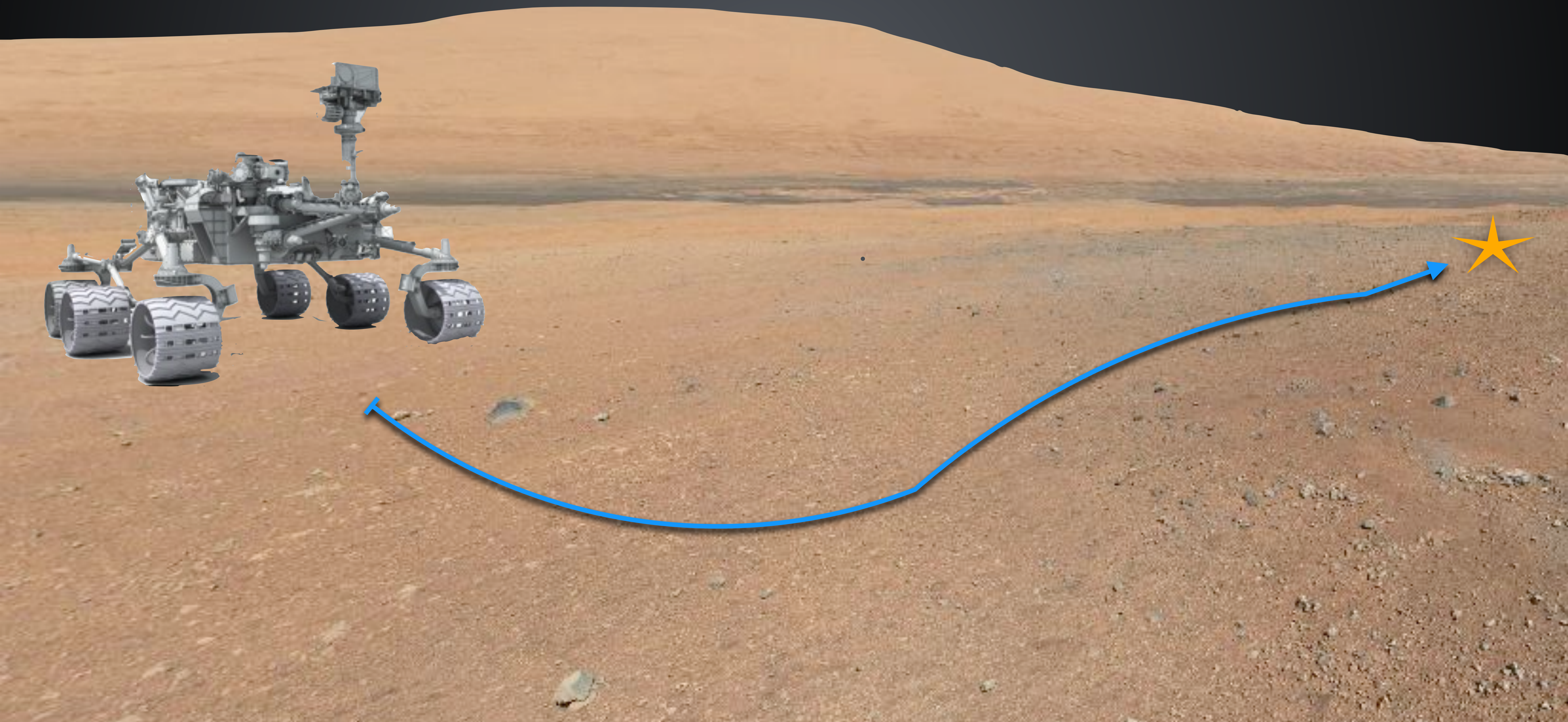
[†] Space and Terrestrial Autonomous Robotic Systems Lab

STARS

[*] Autonomous Space Robotics Lab

ASRL

Visual Odometry: What is it?



Visual Odometry: What is it?

- *Wheel Odometry* can be used to track distance
- In high slip or bumpy terrain, accuracy is poor
- Solution: Visual Odometry *(Moravec, 1980)*

MER *Opportunity* Tracks | April 7, 2009 | Courtesy *JPL*

Visual Odometry: What is it?

The use of image data to estimate the change in *pose*

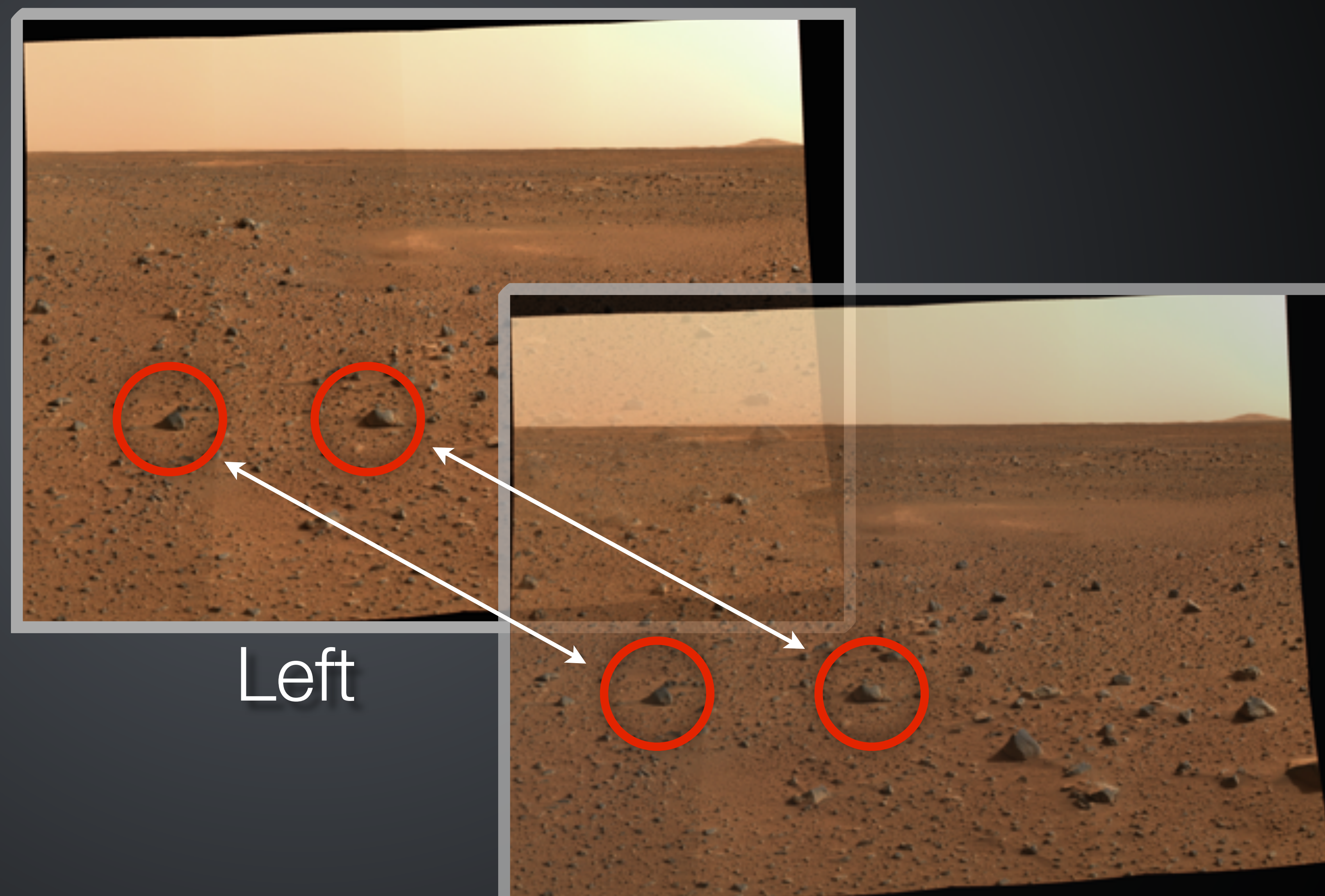
Image Capture
↓
Landmark Detection
↓
Stereo Matching
↓
Landmark Tracking
↓
Outlier Rejection
↓
Motion Solution

Stereo VO
Pipeline

Visual Odometry: What is it?

The use of image data to estimate the change in *pose*

- Image Capture
- Landmark Detection
- Stereo Matching
- Landmark Tracking
- Outlier Rejection
- Motion Solution



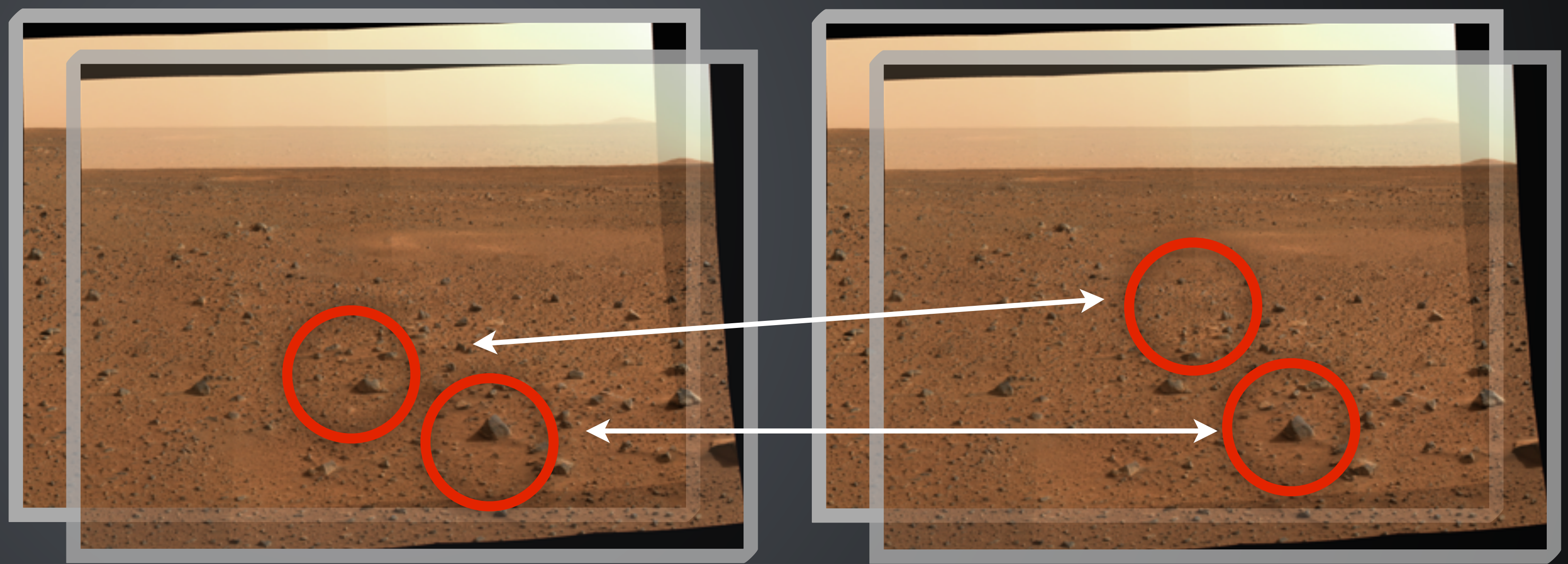
Left

Right

Visual Odometry: What is it?

The use of image data to estimate the change in *pose*

Image Capture
↓
Landmark Detection
↓
Stereo Matching
↓
Landmark Tracking
↓
Outlier Rejection
↓
Motion Solution

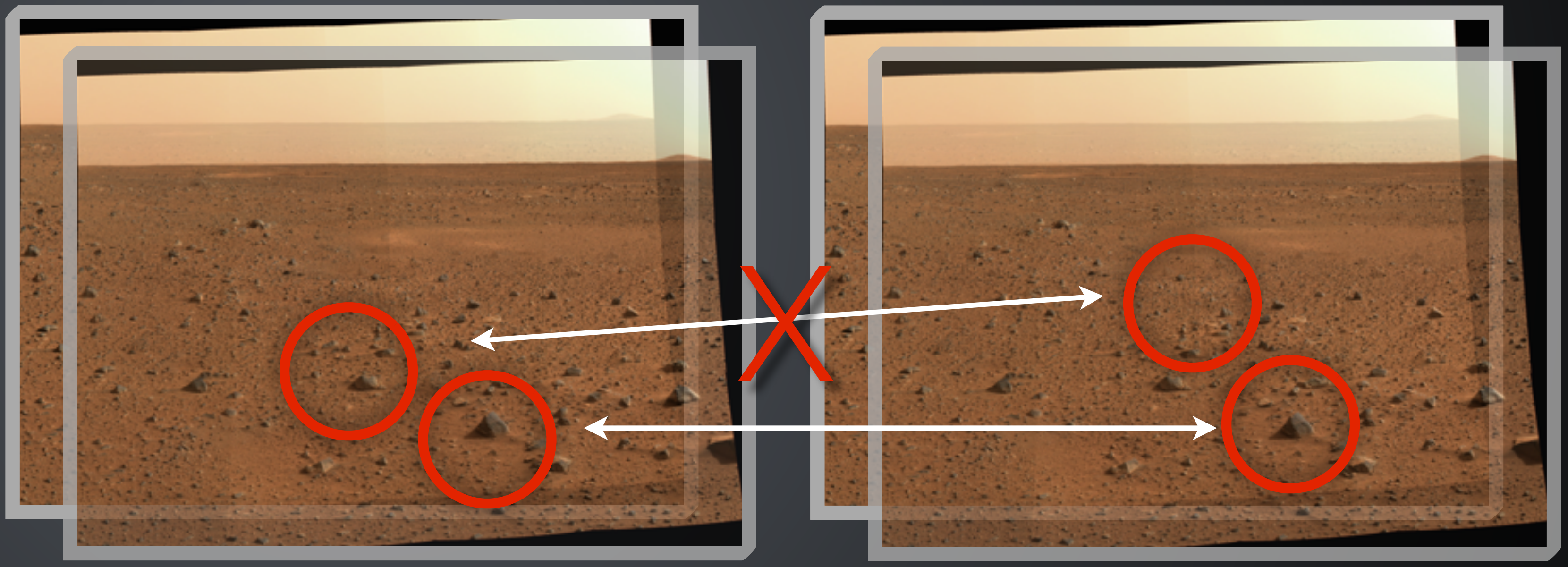


Step i Step i +1

Visual Odometry: What is it?

The use of image data to estimate the change in *pose*

Image Capture
↓
Landmark Detection
↓
Stereo Matching
↓
Landmark Tracking
↓
Outlier Rejection
↓
Motion Solution



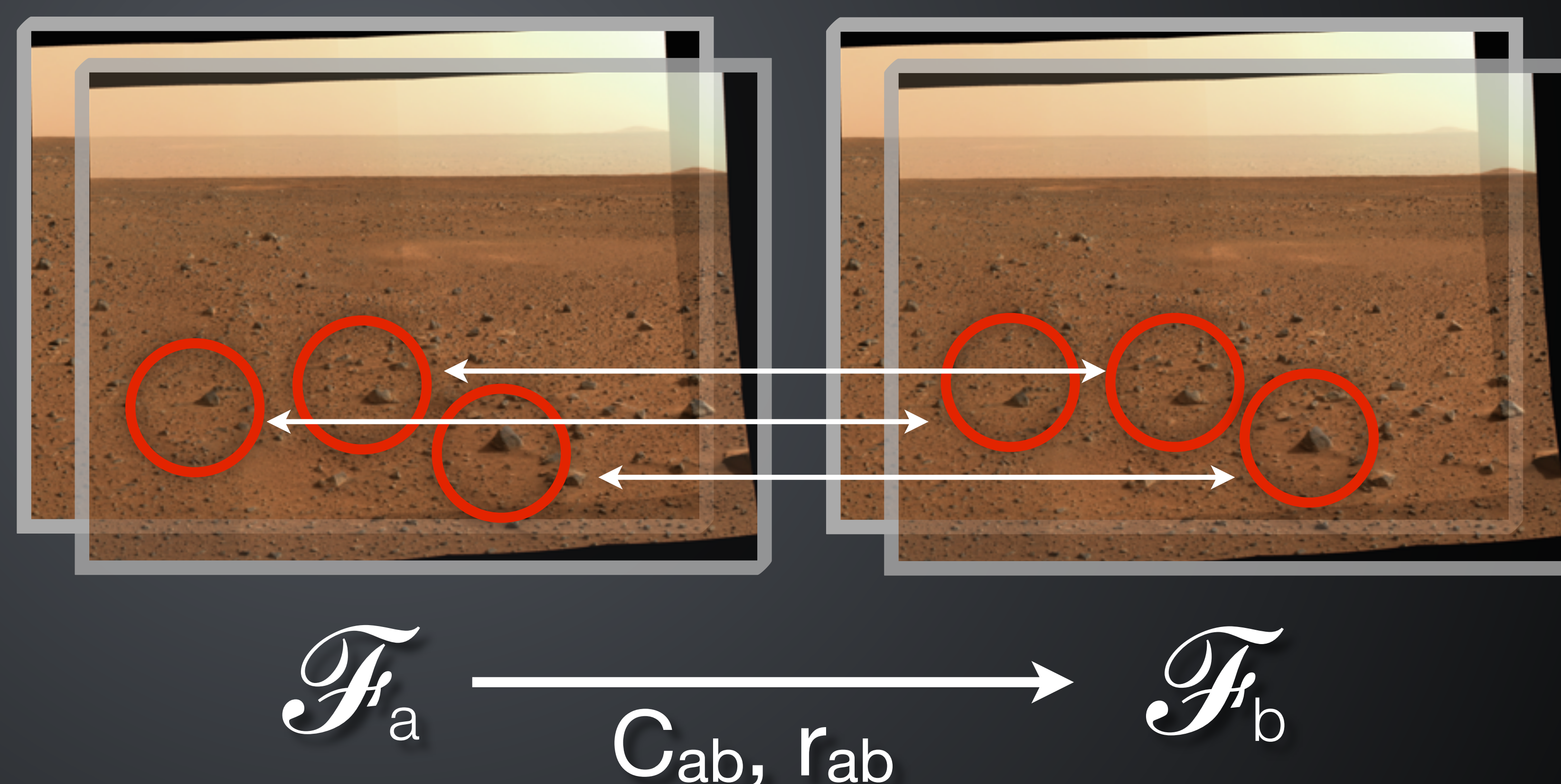
Step i

Step i +1

Visual Odometry: What is it?

The use of image data to estimate the change in *pose*

Image Capture
↓
Landmark Detection
↓
Stereo Matching
↓
Landmark Tracking
↓
Outlier Rejection
↓
Motion Solution



How can we improve VO?

Visual Odometry Improvements

1 Additional Sensors

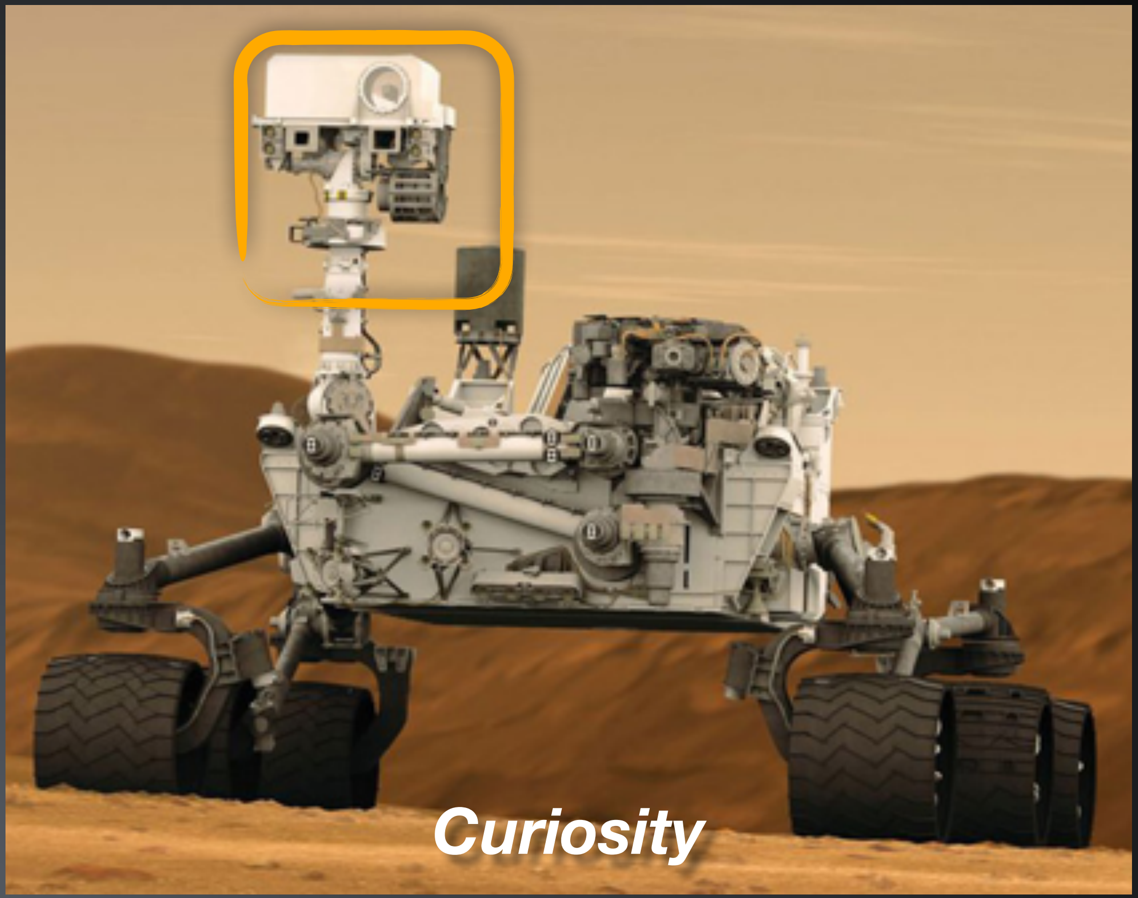
Olson et al. [1]: Orientation sensor leads to error growth reduction (super linear to linear)

2 Additional Constraints

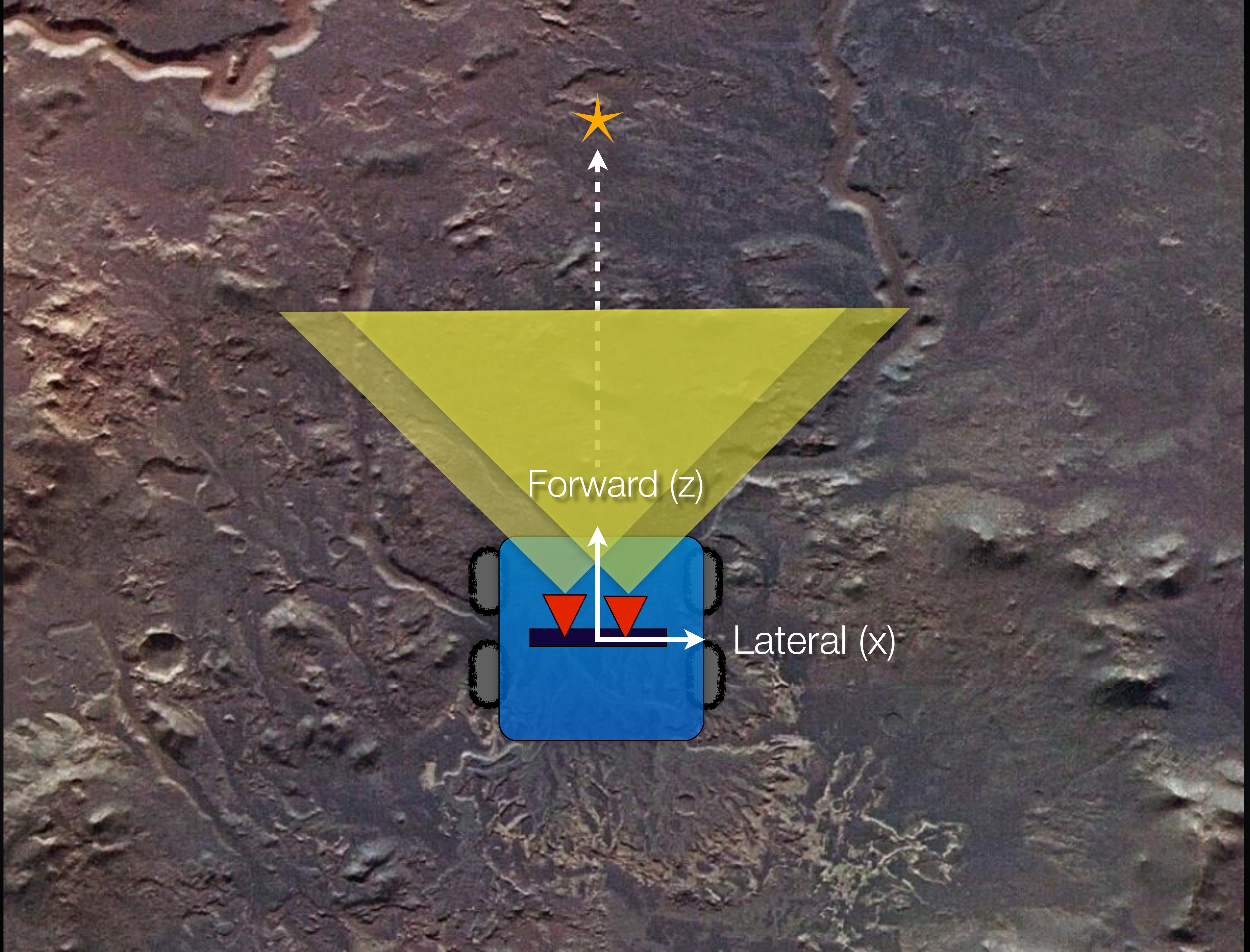
Konolige et al. [2]: Sliding window bundle adjustment (incorporates vision data from multiple poses to improve estimates)

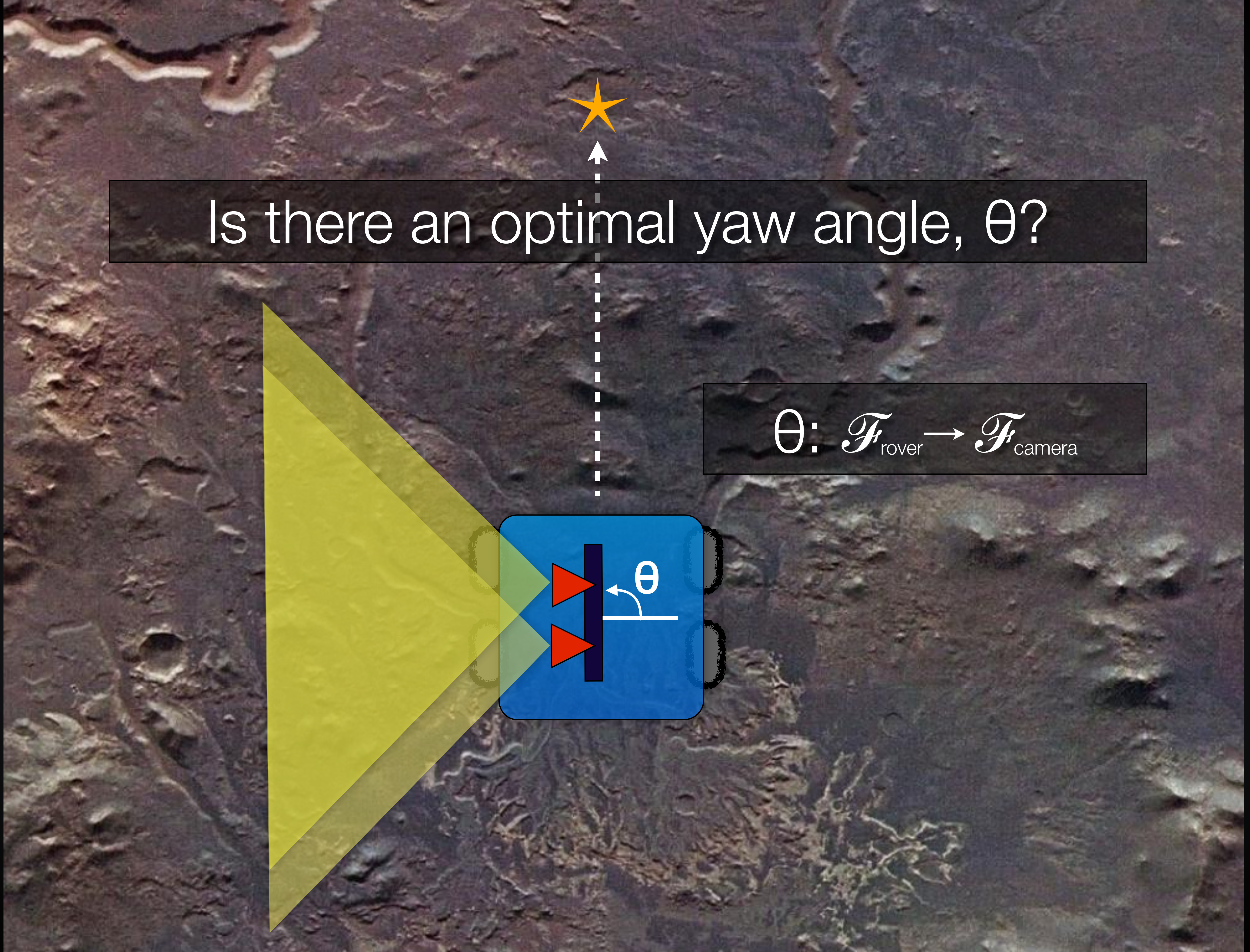
3 Parametric Optimization

Olson et al. [1]: Field of View. *Howard [3]:* Camera Resolution
Kelly and Sukhatme [4]: Camera pitch.



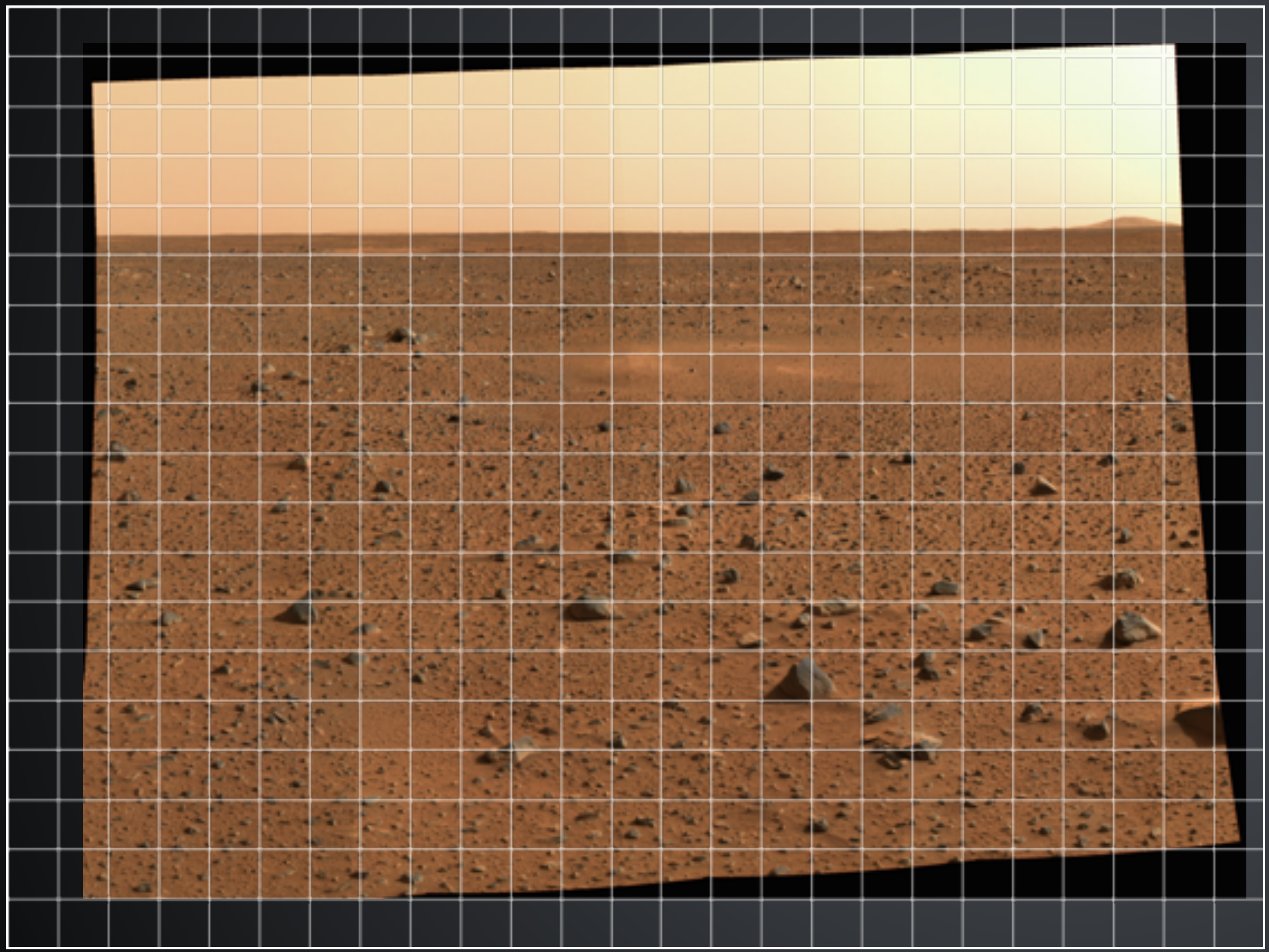
Common Paradigm: Cameras are directed forward.

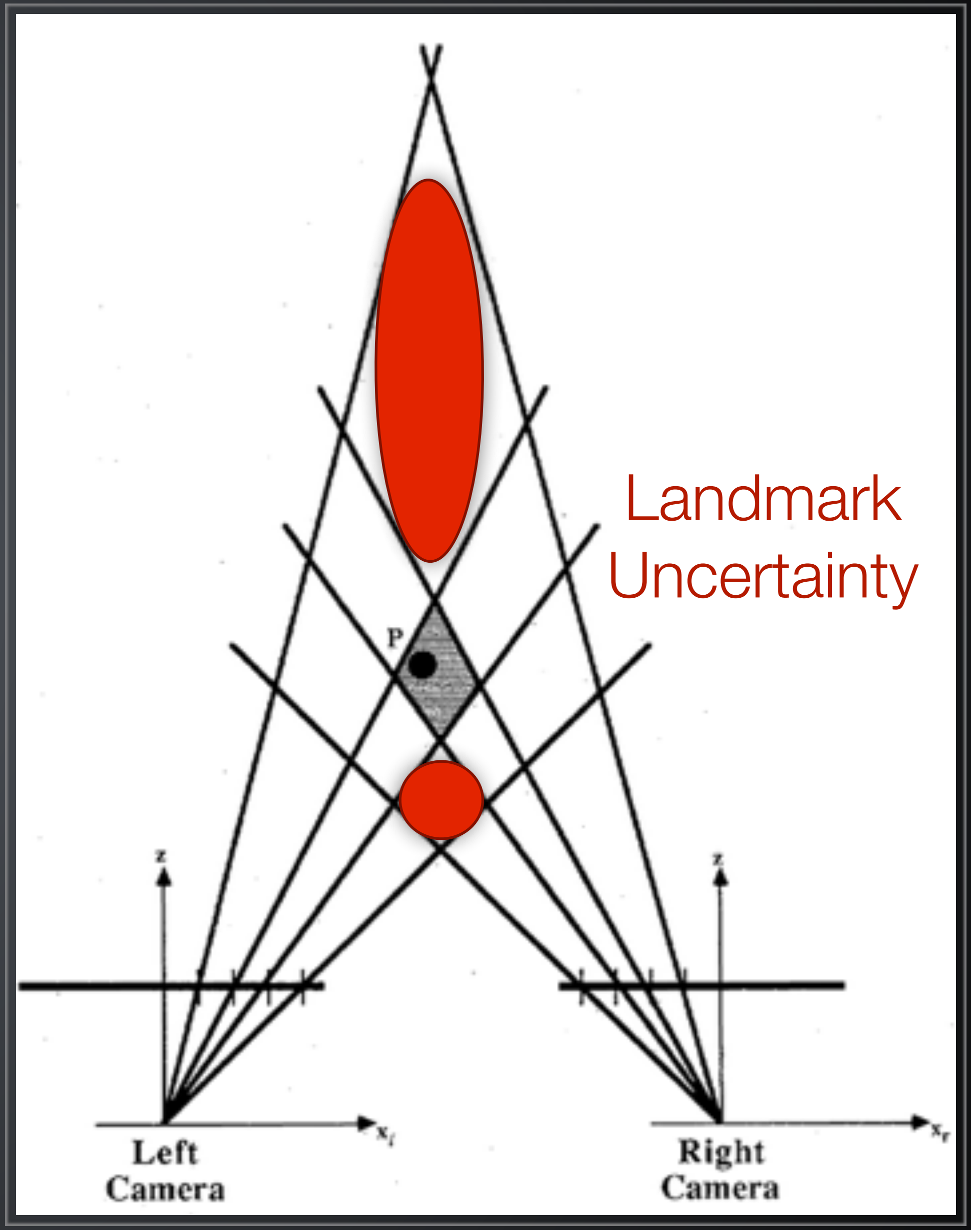

Forward (z)
Lateral (x)


Is there an optimal yaw angle, θ?
$\theta: \mathscr{F}_{rover} \rightarrow \mathscr{F}_{camera}$
θ

Why should changing camera yaw angle have any effect on VO accuracy?

Discrete Image Plane

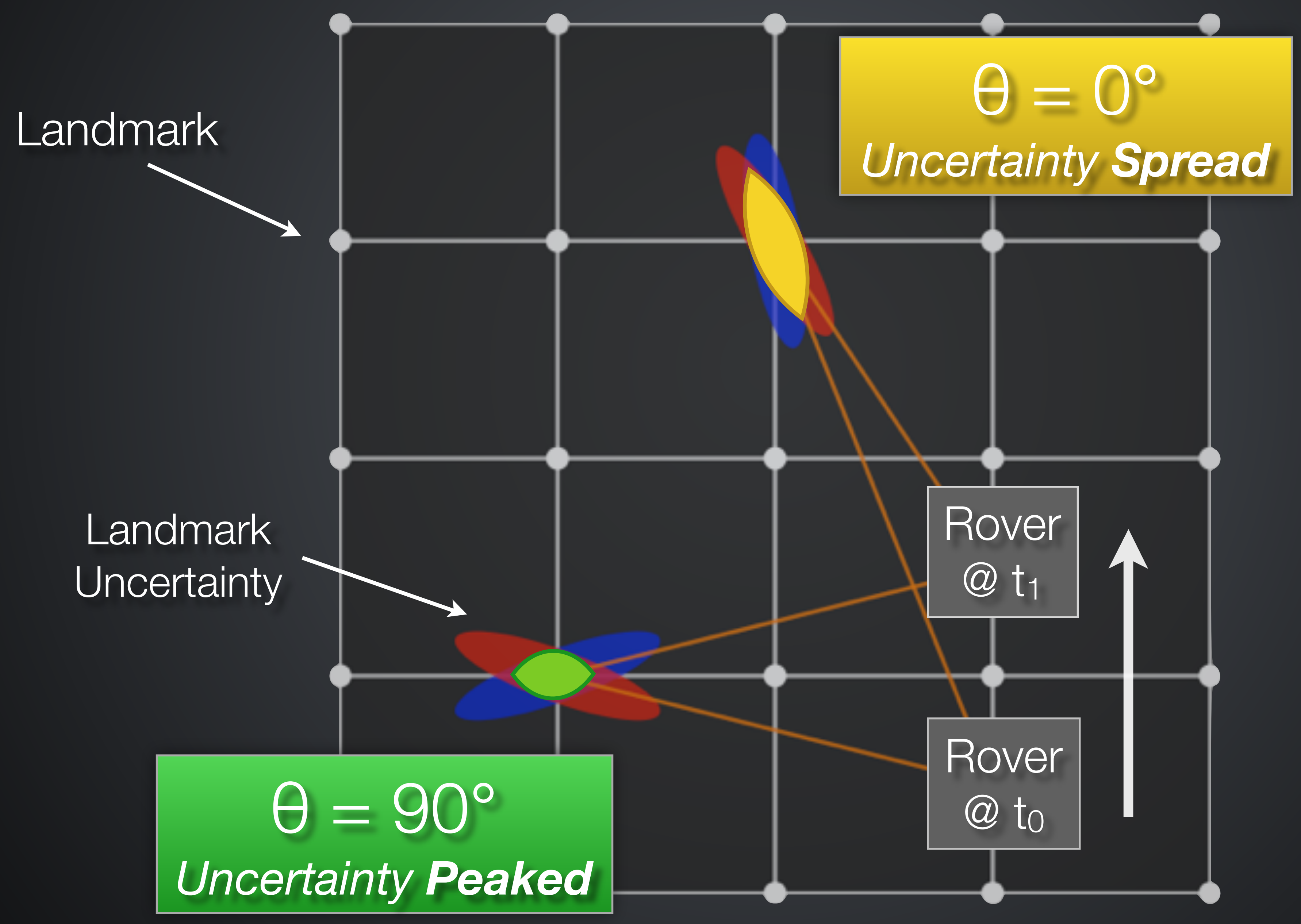




Matthies & Shafer (1987) [5]

Landmark uncertainty grows and becomes elliptical for distant landmarks.

Discrete Image Plane: Consequences



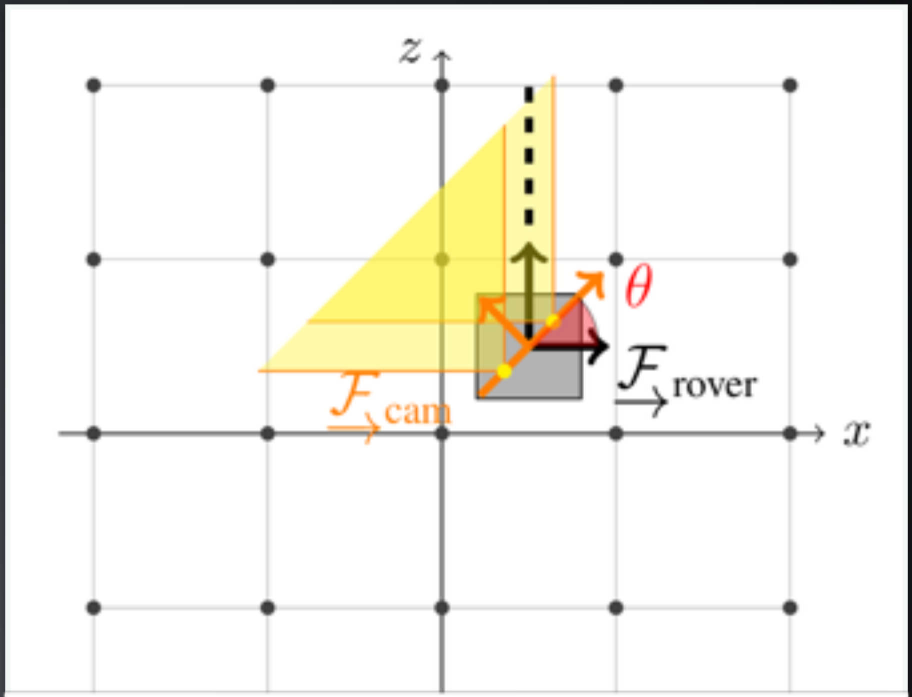
Our **Hypothesis**

For ground vehicles, orienting a stereo camera transverse to the direction of motion will improve VO estimates.

Simulations

Simulation Details

- 2D world with point landmarks
- Pixel measurements corrupted with Gaussian noise
- Landmarks are tracked and matched ideally
- Motion solution: SVD then matrix weighted Gauss-Newton optimization



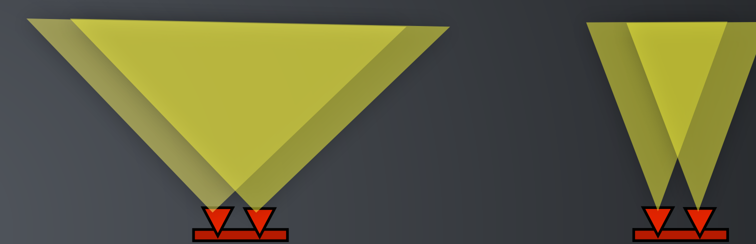
MER: ***Opportunity*** & ***Spirit***

Motion solution algorithm based on one used by the Mars Exploration Rovers (MERs). Detailed in Maimone (2007), Olson (2003).

Variables Investigated

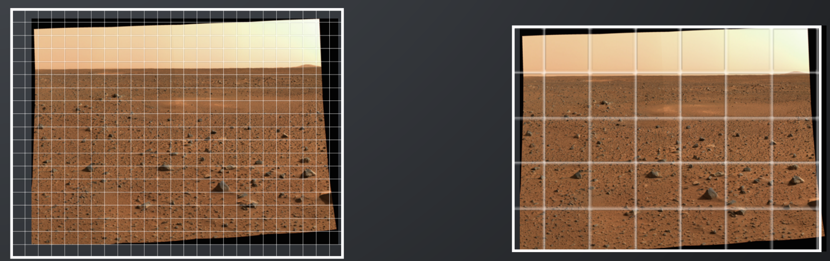
1. Landmark Distribution

2. Field of View



3. Step Size

4. Resolution
(Noise Variance)



Simulation Results

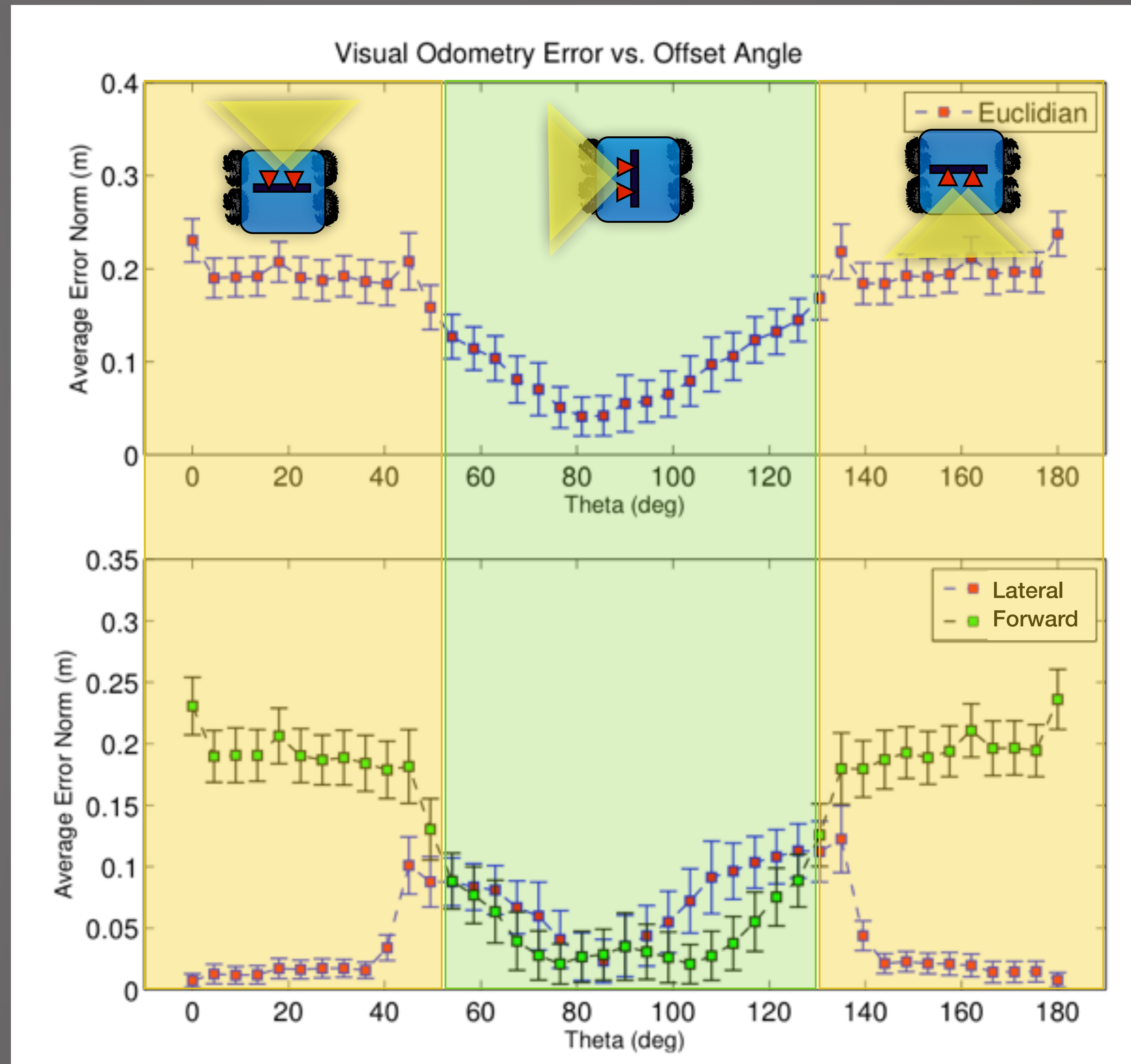
Straight Path through Uniform Distribution

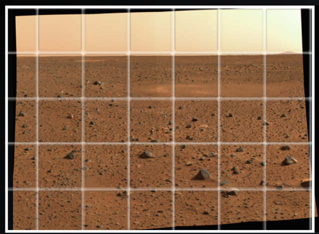
1m straight path through a uniform distribution of 400 landmarks.

500 trials are averaged at each yaw angle, θ.

Pixel noise is set to a variance of 2.25 pixels^2.

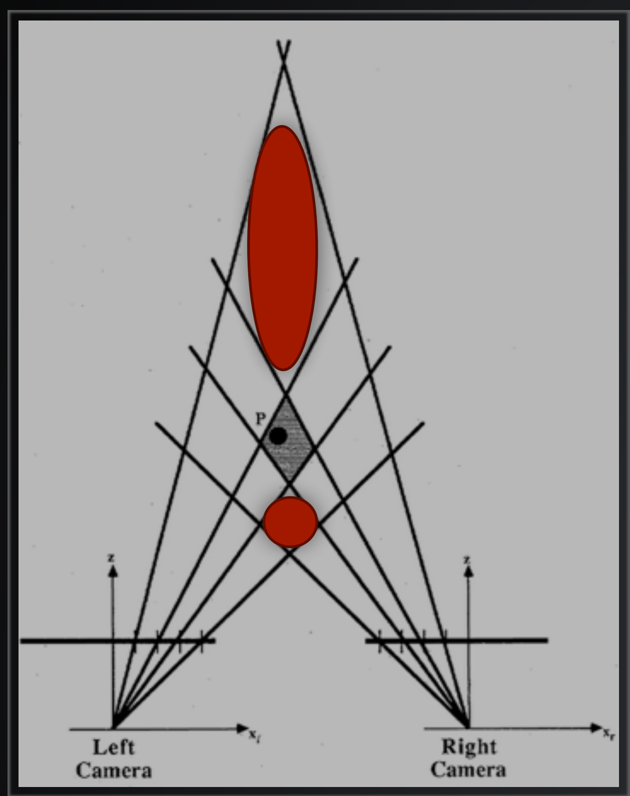
82% reduction in mean Euclidian error at $85 < \theta < 90°$



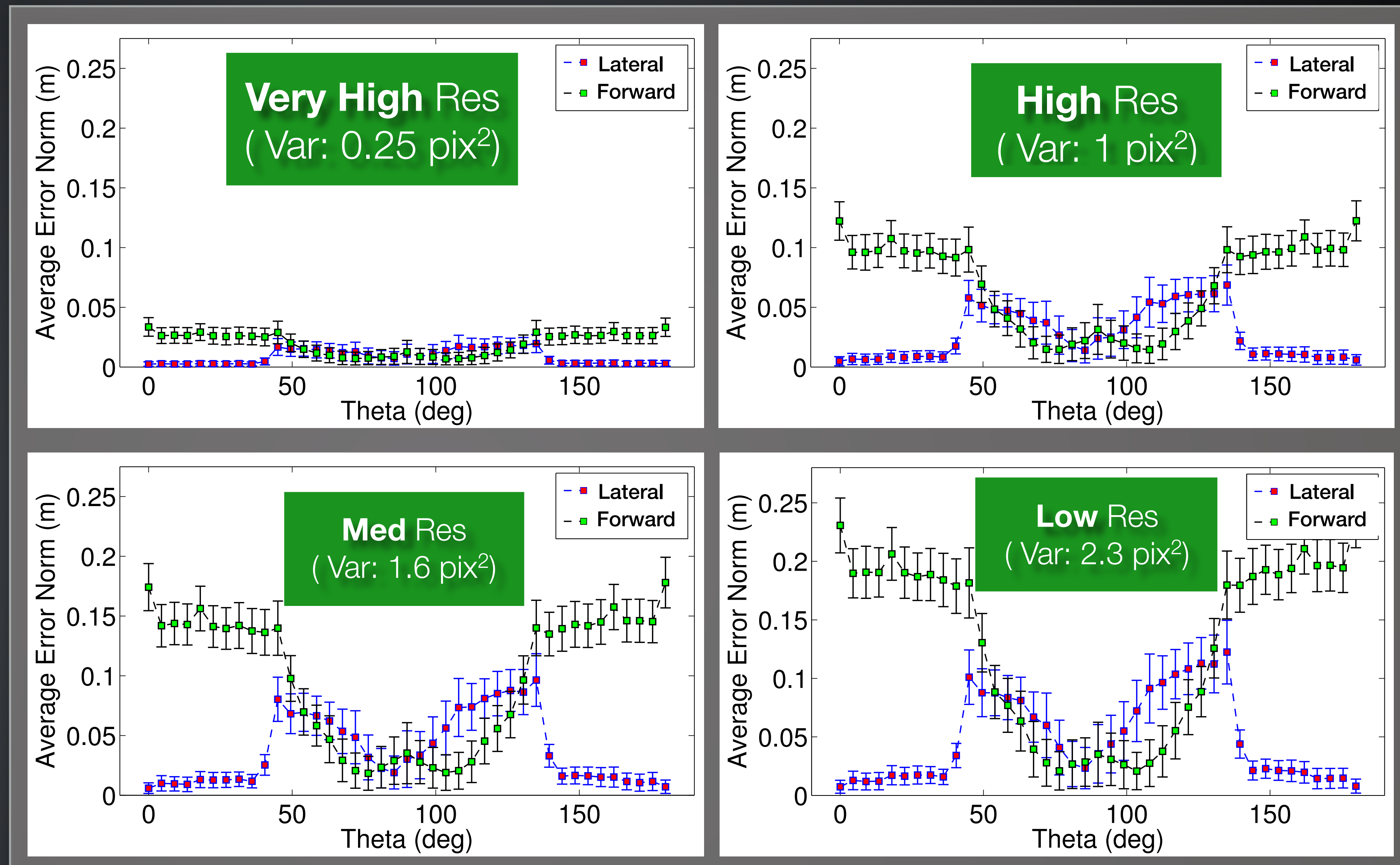


Simulation Results (II)

Effect of Camera Resolution



Effect of image discretization.

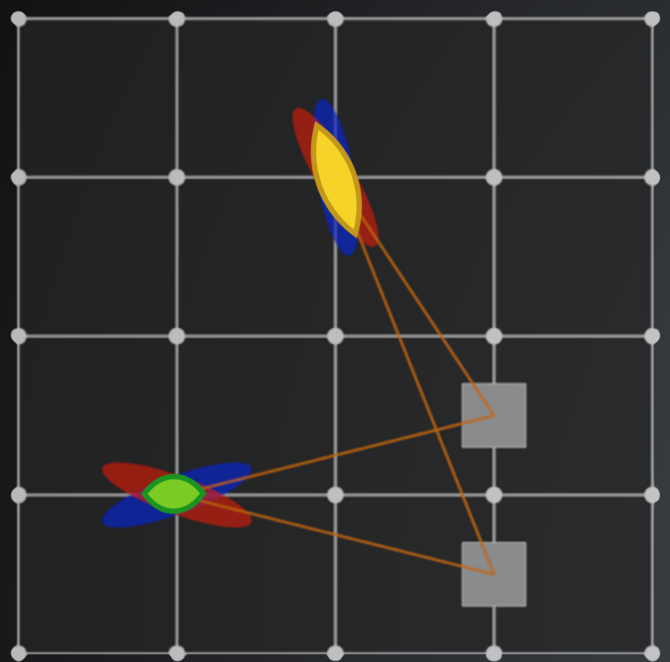


The effect of camera yaw is reduced at high resolutions.

Camera resolution was modelled through additive Gaussian pixel noise of a specified variance.

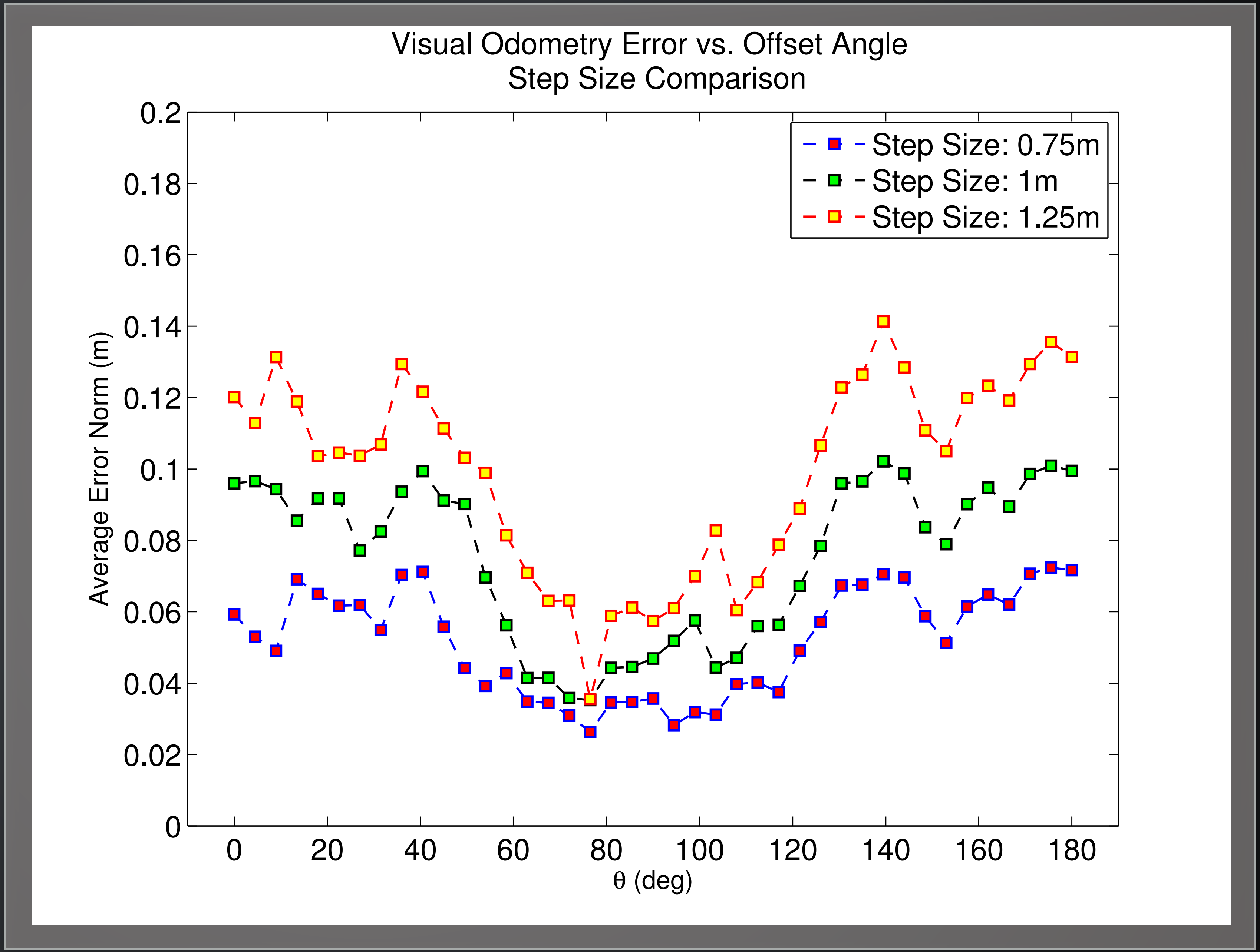
Simulation Results (III)

Random landmark distribution and step size.



A larger step size should reduce the green overlap.

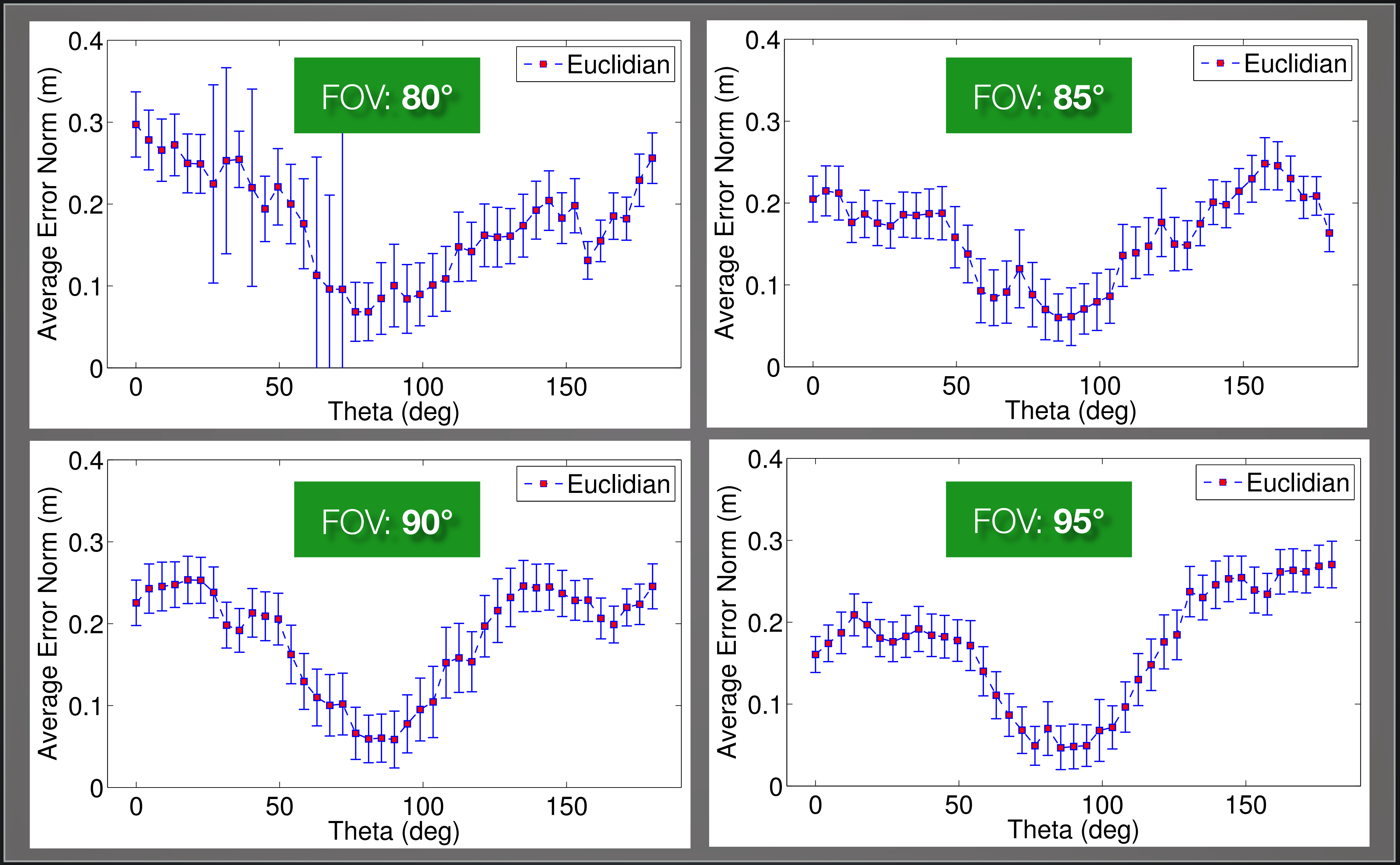
Larger step sizes make the effect of camera yaw more significant.



A random landmark distribution leads to less symmetric results that still maintain a similar reduction at 90°.

Simulation Results (IV)

Effect of field of view.



The optimal field of view is heavily dependent on step size and landmark distribution.

Different field of views were analyzed in a randomized landmark distribution.

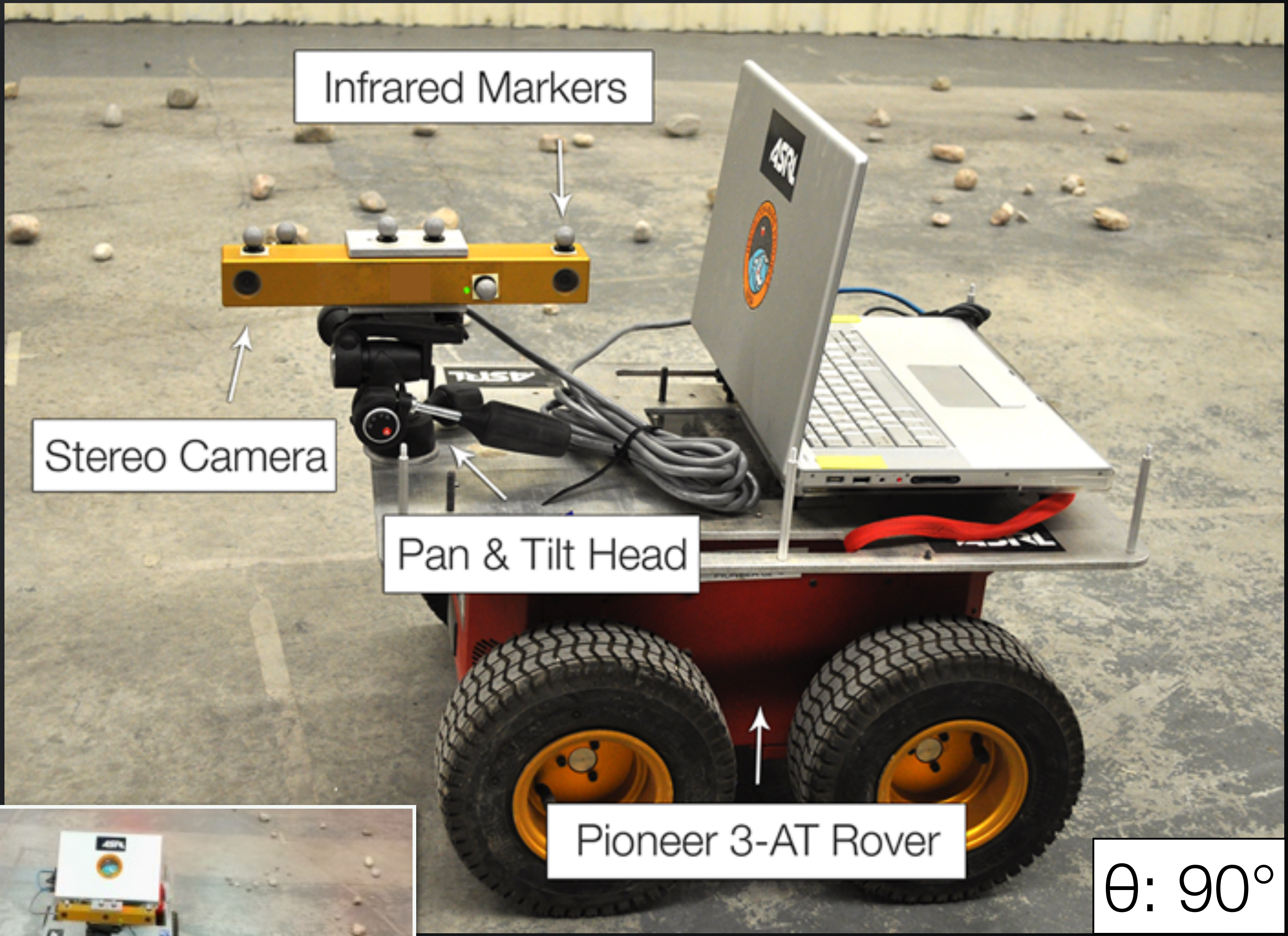
Experiments

Experimental **Setup**

Pioneer 3-AT equipped with a Point Grey Bumblebee XB3 stereo camera on a pan-tilt head.

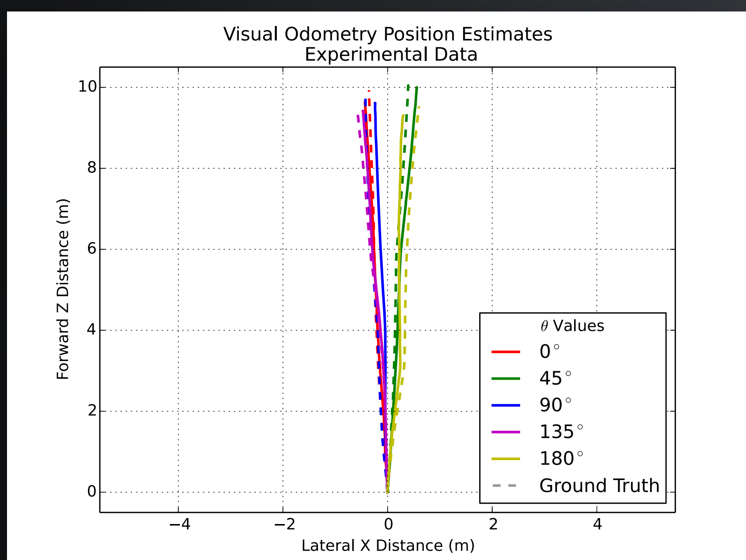
Frame to frame bundle adjustment framework.

VICON Infrared cameras for ground-truth.



Experimental **Results**

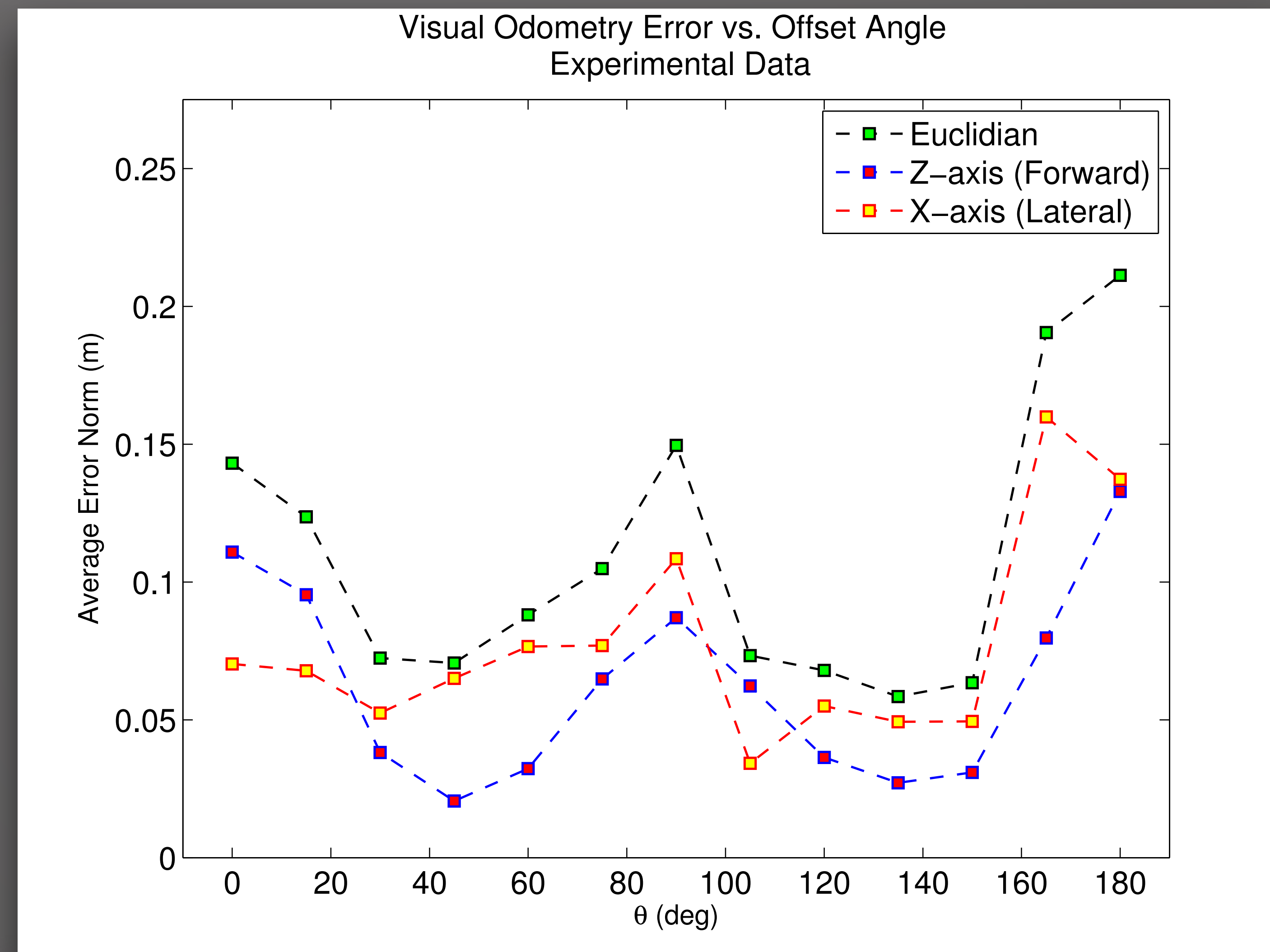
Rover was driven forward for ~10 m.

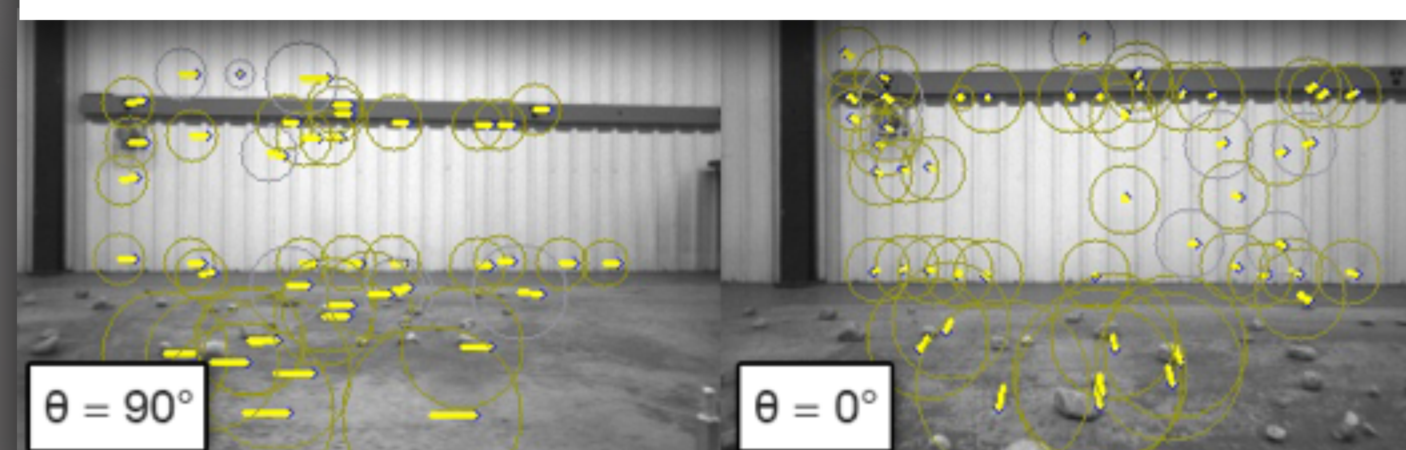


5 trials were averaged for every stereo camera yaw angle (15° intervals).

Total Distance: 650m

Reductions of up to 59% in mean Euclidian error are found at 45° and 135°.

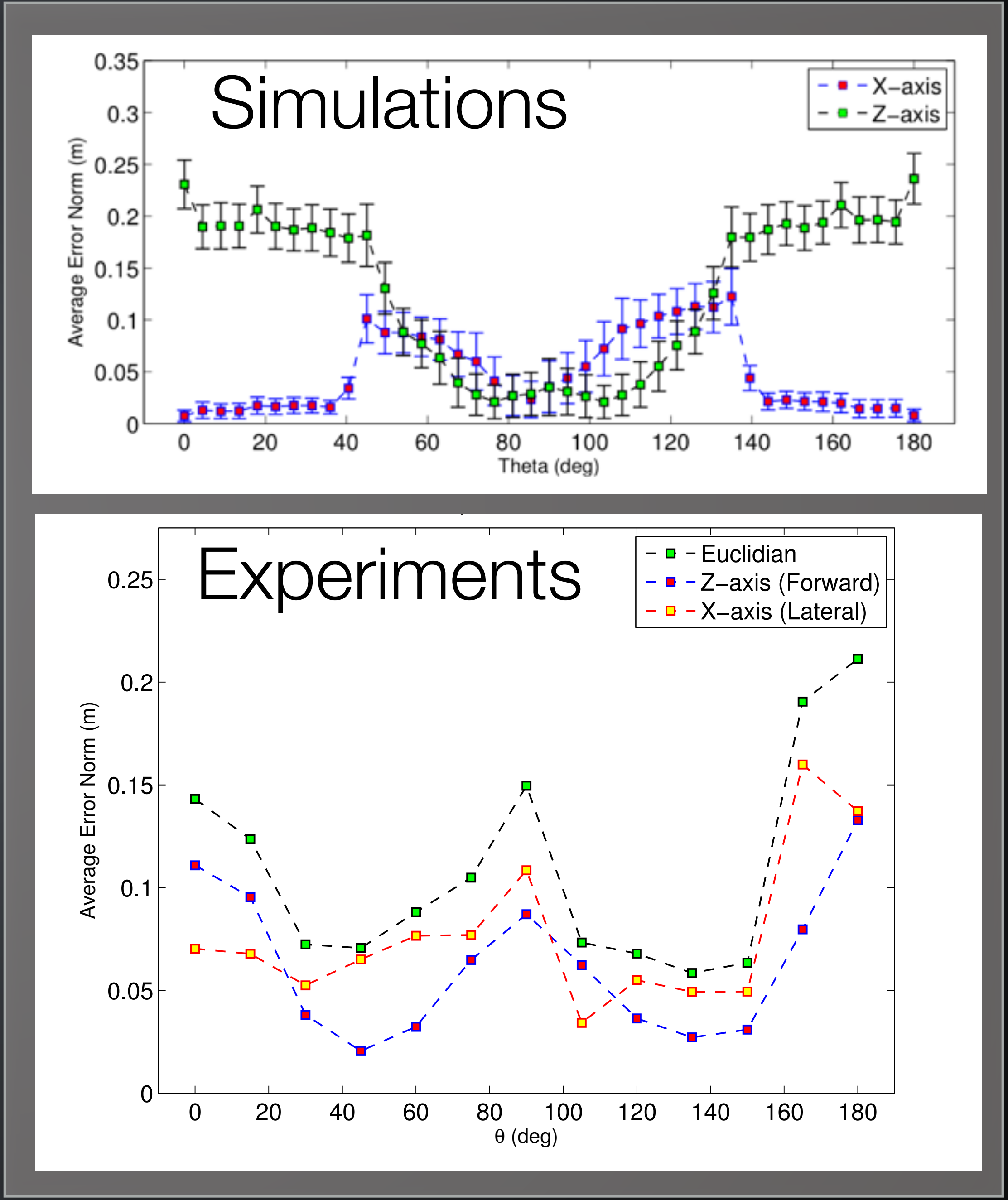




Discussion

Simulations vs. Experiments

- Forward error followed a very similar trend to simulations
- Lateral error was more erratic
- The reverse orientation was consistently worse than looking forward
- Lab geometry, feature distribution, field of view and frame rate differences may explain discrepancies



Conclusions

- By changing the yaw angle of a stereo camera, VO estimation error was reduced by as up to **82%** in simulations and **59%** in real world tests.
- **Computationally limited platforms** (i.e. with a *reduced resolution* and *slower frame rates*) may benefit from a relatively simple configuration change.

Future Work

- Potential future work will include developing an actuator that can change the orientation of the stereo camera based on the rover state to improve VO.

Thanks for listening!

Questions?

Bibliography

[1] C. Olson, L. Matthies, M. Schoppers, and M. Maimone, “Rover navigation using stereo ego-motion,” Robotics and Autonomous Systems, vol. 43, pp. 215–229, 2003.

[2] K. Konolige, M. Agrawal, and J. Sola, “Large-scale visual odometry for rough terrain,” in Robotics Research. Springer, 2011, pp. 201–212.

[3] A. Howard, “Real-time stereo visual odometry for autonomous ground vehicles,” in Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on. IEEE, 2008, pp. 3946–3952.

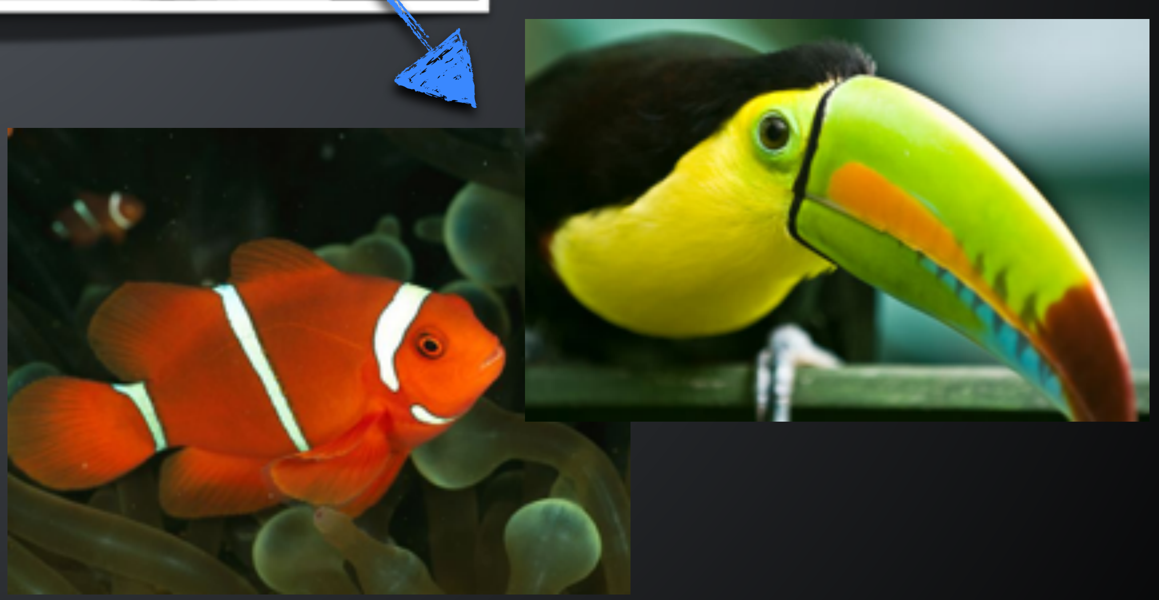
[4] J. Kelly and G. S. Sukhatme, “An experimental study of aerial stereo visual odometry,” in Proc. 6th IFAC Symp. Intelligent Autonomous Vehicles (IAV’07), Toulouse, France, Sep 2007.

[5] L. Matthies and S. A. Shafer, “Error modelling in stereo navigation,” *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 3, 1987.

Additional Slides

Visual **Odometry**: Biomimicry

Human Binocular Vision



Vision in Nature

The nitty gritty **mathematics**

Stereo Camera Model

$$\mathbf{p} = \mathbf{g}(\mathbf{y}) = \begin{bmatrix} \frac{b}{2}\frac{(u_l+u_r)}{u_l-u_r} \\ \frac{b}{2}\frac{(v_l+v_r)}{u_l-u_r} \\ \frac{bf}{u_l-u_r} \end{bmatrix}$$

Matrix Weighted, Gauss-Newton Optimization

$$\mathcal{L} = \frac{1}{2}\sum_{j=1}^{N}\left(\mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba})\right)^T \boldsymbol{\Gamma}^j \left(\mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba})\right)$$

$$\boldsymbol{\Gamma}^j = \left(\mathbf{G}_b^j \mathbf{R}_b^j \mathbf{G}_b^{j^T} + \mathbf{C}_{ba}\mathbf{G}_a^j \mathbf{R}_a^{j^T} \mathbf{G}_a^{j^T} \mathbf{C}_{ba}^T\right)^{-1}$$

$$\mathbf{G}_b^j = \left.\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right|_{\mathbf{f}(\mathbf{p}_b^j)} \qquad \mathbf{G}_a^j = \left.\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right|_{\mathbf{f}(\mathbf{p}_a^j)}$$

Scalar Weighted, Analytic Initial Estimate

$$\mathcal{J} = \frac{1}{2}\sum_{j=1}^{N} w^j \left(\mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba})\right)^T \left(\mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba})\right)$$

$$w^j = \left[\det\left(\boldsymbol{\Sigma}_a^j\right) + \det\left(\boldsymbol{\Sigma}_b^j\right)\right]^{-1}, \quad w = \sum_{j=1}^{N} w^j$$

$$\boldsymbol{\Sigma}^j = \left[\left.\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right|_{\mathbf{y}^j}\right] \begin{bmatrix} \mathbf{R}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_r \end{bmatrix} \left[\left.\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right|_{\mathbf{y}^j}\right]^T$$

Presentation **Outline**

1. Stereo Visual Odometry
   - *Dead reckoning through vision*
   - *Ways to improve VO*
2. Our Hypothesis
   - *Uncertainty from discretization leads to non-uniform localization accuracy*
3. Simulations
4. Experiments
5. Conclusions and Future Work