

# Optimizing Camera Perspective for Stereo Visual Odometry

Valentin Peretroukhin, Jonathan Kelly and Timothy D. Barfoot

*Institute for Aerospace Studies*

*University of Toronto*

*Toronto, Canada*

{v.peretroukhin, js.kelly, tim.barfoot}@utoronto.ca

**Abstract**—Visual Odometry (VO) is an integral part of many navigation techniques in mobile robotics. In this work, we investigate how the orientation of the camera affects the overall position estimates recovered from stereo VO. Through simulations and experimental work, we demonstrate that this error can be significantly reduced by changing the perspective of the stereo camera in relation to the moving platform. Specifically, we show that orienting the camera at an oblique angle to the direction of travel can reduce VO error by up to 82% in simulations and up to 59% in experimental data. A variety of parameters are investigated for their effects on this trend including frequency of captured images and camera resolution.

**Keywords**—visual odometry; mobile robotics; stereo vision; localization

## I. INTRODUCTION

The ability to navigate within novel environments is an important part of many autonomous robotics applications. In many terrestrial and extraterrestrial settings, this task is made significantly more difficult by the lack of an external positioning system (such as GPS) to provide an accurate estimate of a robot’s egomotion. Traditional odometry techniques that use wheel encoders often provide poor approximations to true motion over any substantial distance due to wheel slip. In this paper, we study a technique called *visual odometry* (VO) [1]–[3] that aims to provide accurate localization estimates through the use of an imaging sensor. We focus on a classic implementation of VO, first pioneered by Hans Moravec [4], that uses a stereo camera. As with all dead-reckoning techniques, stereo VO exhibits error that grows without bound with the distance travelled. Unlike other techniques, however, stereo VO can produce accurate pose estimates over significant distances, with errors as low as a few percent of the distance travelled on trajectories of several hundred meters [5]. It is also particularly useful because, unlike monocular implementations, the absolute scale of the motion can be recovered without the use of any additional sensors or prior knowledge of scene structure.

One aspect of stereo visual odometry has been seldom investigated: the effect of the orientation of the stereo camera on motion estimates. To the best of our knowledge, nearly all implementations in the literature point the stereo camera forwards, in line with the motion of the rover [1]–[3]. The motivation behind this stems from the fact

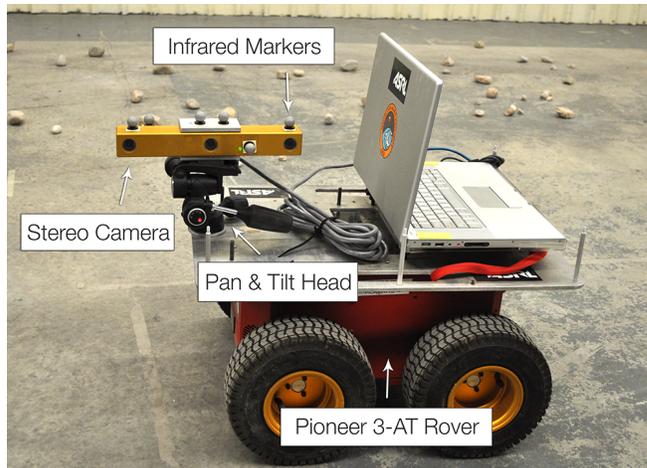


Figure 1: The experimental platform: a stereo camera mounted on a Pioneer 3-AT rover base.

that, with a forward-facing camera, collected images can be used for other tasks such as collision avoidance and path planning. Many modern robotics platforms, however, incorporate multiple vision sensors, allowing one stereo camera to be dedicated solely to visual odometry.

With this in mind, we seek to determine which camera orientation minimizes error in stereo VO motion estimates. Our work focuses on a simplified rover configuration in which the camera can be rotated about a vertical axis with respect to the chassis and fixed prior to the start of traversal. We present both simulated and experimental data (from the Pioneer rover depicted in Figure 1) which show that the error in position estimates can be minimized by directing the camera to one side of the moving vehicle.

To understand why a particular camera orientation may result in improved motion estimates, consider a simplified two-dimensional world presented in Figure 2. As first noted by Matthies and Shafer [6], image discretization introduces uncertainty in the positions of observed landmarks. This uncertainty generally increases for landmarks that are further away from the camera, creating ellipsoidal probability densities that elongate as landmarks become more distant. When the stereo camera is directed perpendicular to the direction of travel, the overlap between uncertainty ellipsoids corre-

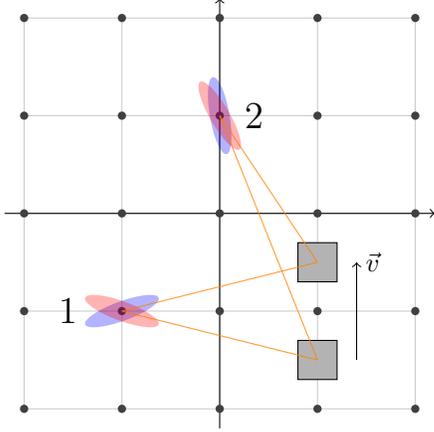


Figure 2: Effect of directing a stereo camera perpendicular to rover motion. Two dimensional Gaussian distributions are shown by red and blue sigma contours. Landmarks (shown as black dots) can be better localized because their positional uncertainty is reduced (i.e., overlapping area of uncertainty distributions is minimized). Landmark 1 can be localized better than landmark 2.

sponding to consecutive measurements of a given landmark decreases. Thus, the combination of measurements should improve landmark localization and, hence, VO accuracy.

## II. RELATED WORK

Visual odometry was first pioneered by Moravec [4], with applications motivated by the NASA Mars exploration program [7]. Moravec outlined the basic 6 steps that have remained the building blocks of many modern VO pipelines [8]. Each of these steps (presented in Figure 3) has been studied extensively, as discussed in [7] and [9].

In Moravec’s approach, observed landmarks are back-projected into 3D space, and the motion between two poses is solved for by aligning the corresponding point clouds using a scalar weighted nonlinear optimization. Several years later, Matthies [15] elaborated on this approach with an analytical method (based on work by Horn [10]) that uses singular value decomposition to efficiently solve the scalar-weighted optimization problem. In a landmark paper, Matthies and Shafer [6] presented a technique using matrix-weighted minimization that models the landmark locations with 3D Gaussian distributions to better account for error from image discretization. The final transformation was then equivalent to the *maximum likelihood state estimate* given the observed landmark locations.

Modern visual odometry algorithms based on these two early research efforts have been implemented on the NASA Mars Exploration Rovers [2], and for long distance terrestrial navigation [1]. Various attempts at improving position estimates recovered from VO have been developed. Olson et al. [3] noted that the super-linear error growth of visual

odometry estimates could be improved to grow linearly by incorporating an orientation sensor. Konolige et al. [11] developed a *sliding window bundle adjustment* technique that incorporates constraints from features tracked over multiple robot poses to reduce drift.

Various parametric optimizations have also been investigated. Olson et al. [3] showed through simulations that a field-of-view value of  $35^\circ$  provided an optimal balance between angular resolution and landmark tracking. Howard [12] found that VO RMS error decreased by nearly 90% after increasing camera resolution from  $160 \times 120$  pixels to  $640 \times 480$  pixels.

To the best of our knowledge, the only work that has examined the effect of camera orientation on stereo visual odometry estimates is that of Kelly and Sukhatme [5]. The authors presented an experimental study conducted using a radio-controlled helicopter, with two different forward-facing camera orientations. They concluded that orienting the camera  $66^\circ$  below the horizon produced lower pose estimation errors than orienting it at a direction close to nadir,  $85^\circ$  below the horizon.

## III. VISUAL ODOMETRY PIPELINE

To determine the effect of rotating the stereo camera with respect to the rover frame, a 2D stereo visual odometry simulation was built. Landmarks are taken to be points in space projected onto a perfectly rectified pinhole stereo camera. Stereo matching is assumed to be known. The detected image coordinates are corrupted with zero-mean Gaussian noise. The remainder of the visual odometry pipeline largely resembles that presented by Maimone et al. [2]. Steps 4 through 6 (see Figure 3) of the pipeline are outlined here.

We first define two frames,  $\mathcal{F}_a$  and  $\mathcal{F}_b$ , that represent the camera pose at two subsequent time steps. The coordinates of a 3D point observed in  $\mathcal{F}_a$ ,  $\mathbf{p}_a$ , can be transformed into its coordinates in  $\mathcal{F}_b$ ,  $\mathbf{p}_b$ , as follows:

$$\mathbf{p}_b = \mathbf{C}_{ba}(\mathbf{p}_a - \mathbf{r}_a^{ba}), \quad (1)$$

where  $\mathbf{r}_a^{ba}$  are the coordinates of the origin of  $\mathcal{F}_b$  expressed in  $\mathcal{F}_a$  and  $\mathbf{C}_{ba}$  is the rotation matrix from  $\mathcal{F}_a$  to  $\mathcal{F}_b$ .

Our stereo camera model projects points from the 3D camera frame (with an origin half way between the two camera pinholes) into 2D image coordinates in the left and right camera images:

$$\mathbf{y} = \mathbf{f}(\mathbf{p}) = \begin{bmatrix} u_l \\ v_l \\ u_r \\ v_r \end{bmatrix} = \frac{1}{p_3} \begin{bmatrix} f(p_1 + \frac{1}{2}b) \\ fp_2 \\ f(p_1 - \frac{1}{2}b) \\ fp_2 \end{bmatrix}, \quad (2)$$

where  $p_i$  are the components of  $\mathbf{p}$  and  $u_l$ ,  $u_r$ ,  $v_l$ ,  $v_r$  are the left and right horizontal and vertical image coordinates, respectively. The focal length of the camera is  $f$  and its baseline is  $b$ .



Figure 3: Stereo Visual Odometry Pipeline.

Given a set of pixel coordinates, the 3D location of a point in the stereo camera frame can be found using the inverse of our camera model,  $\mathbf{g}$ :

$$\mathbf{p} = \mathbf{g}(\mathbf{y}) = \begin{bmatrix} \frac{b}{2} \frac{(u_l + u_r)}{u_l - u_r} \\ \frac{b}{2} \frac{(v_l + v_r)}{u_l - u_r} \\ \frac{bf}{u_l - u_r} \end{bmatrix}. \quad (3)$$

#### A. Scalar Weighted Motion Solution

Assuming all landmarks are tracked correctly, our goal is to calculate the transformation between  $\mathcal{F}_a$  and  $\mathcal{F}_b$ , parametrized by  $\mathbf{C}_{ba}$  and  $\mathbf{r}_a^{ba}$ . We proceed by minimizing a sum of squared errors objective function, weighted by scalar weights  $w^j$ :

$$\mathcal{J} = \frac{1}{2} \sum_{j=1}^N w^j \left( \mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba}) \right)^T \left( \mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba}) \right). \quad (4)$$

This is done by defining

$$\mathbf{W}_{ba} = \frac{1}{w} \sum_{j=1}^N w^j \left( \mathbf{p}_b^j - \mathbf{u}_b \right) \left( \mathbf{p}_a^j - \mathbf{u}_a \right)^T, \quad (5)$$

where

$$w^j = \left[ \det(\Sigma_a^j) + \det(\Sigma_b^j) \right]^{-1}, \quad w = \sum_{j=1}^N w^j, \quad (6)$$

$$\mathbf{u}_b = \sum_{j=1}^N w^j \mathbf{p}_b^j, \quad \mathbf{u}_a = \sum_{j=1}^N w^j \mathbf{p}_a^j, \quad (7)$$

$$\Sigma^j = \begin{bmatrix} \frac{\partial \mathbf{g}}{\partial \mathbf{y}} \Big|_{\mathbf{y}^j} \end{bmatrix} \begin{bmatrix} \mathbf{R}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_r \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{g}}{\partial \mathbf{y}} \Big|_{\mathbf{y}^j} \end{bmatrix}^T. \quad (8)$$

The covariance of each pixel measurement,  $\mathbf{R}_l$  and  $\mathbf{R}_r$ , is taken to be  $\mathbf{I}$ .

Singular value decomposition of  $\mathbf{W}_{ba}$  leads to

$$\mathbf{W}_{ba} = \mathbf{V} \mathbf{S} \mathbf{U}^T. \quad (9)$$

Based on [13], the rotation,  $\mathbf{C}_{ba}$ , and the translation  $\mathbf{r}_a^{ba}$  can then be extracted as follows:

$$\mathbf{C}_{ba} = \mathbf{V} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(\mathbf{U})\det(\mathbf{V}) \end{bmatrix} \mathbf{U}^T \quad (10)$$

$$\mathbf{r}_a^{ba} = -\mathbf{C}_{ba}^T \mathbf{u}_b + \mathbf{u}_a.$$

This analytical solution is embedded within an outlier rejection scheme, RANSAC (Random Sample Consensus) [14], to eliminate tracked landmark pairs that do not agree with the majority motion solution.

#### B. Matrix Weighted Motion Solution

The transformation calculated above is then used as an initial guess,  $\{\bar{\mathbf{C}}_{ba}, \bar{\mathbf{r}}_a^{ba}\}$ , in an iterative matrix weighted approach. This method is adapted from [15], and uses the rotation matrix perturbation scheme presented in [16].

The objective function we wish to minimize is

$$\mathcal{L} = \frac{1}{2} \sum_{j=1}^N \left( \mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba}) \right)^T \mathbf{\Gamma}^j \left( \mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba}) \right), \quad (11)$$

with

$$\mathbf{\Gamma}^j = \left( \mathbf{G}_b^j \mathbf{R}_b^j \mathbf{G}_b^{jT} + \mathbf{C}_{ba} \mathbf{G}_a^j \mathbf{R}_a^j \mathbf{G}_a^{jT} \mathbf{C}_{ba}^T \right)^{-1}, \quad (12)$$

$$\mathbf{G}_b^j = \frac{\partial \mathbf{g}}{\partial \mathbf{y}} \Big|_{\mathbf{f}(\mathbf{p}_b^j)}, \quad (13)$$

$$\mathbf{G}_a^j = \frac{\partial \mathbf{g}}{\partial \mathbf{y}} \Big|_{\mathbf{f}(\mathbf{p}_a^j)}, \quad (14)$$

where  $\mathbf{R}_b^j$  and  $\mathbf{R}_a^j$  are covariances of each point in image space, taken to be  $\mathbf{I}$  as before.

We proceed by perturbing the initial estimate with perturbation vector  $\boldsymbol{\xi}$ , composed of a translation vector  $\boldsymbol{\epsilon}$  and an Euler angle parametrization  $\boldsymbol{\phi}$ :

$$\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\epsilon} \\ \boldsymbol{\phi} \end{bmatrix} \quad (15)$$

$$\mathbf{r}_a^{ba} = \bar{\mathbf{r}}_a^{ba} + \boldsymbol{\epsilon}, \quad (16)$$

$$\mathbf{C}^{ba} = e^{-\boldsymbol{\phi}^\times} \bar{\mathbf{C}}^{ba} \approx (\mathbf{I} - \boldsymbol{\phi}^\times) \bar{\mathbf{C}}^{ba}. \quad (17)$$

Here,  $(\cdot)^\times$  is the skew-symmetric cross-product operator:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}^\times := \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}.$$

Inserting this into (11), we arrive at a cost function that is quadratic in the perturbations:

$$\mathcal{L} \approx \frac{1}{2} \sum_{j=1}^N (\bar{\boldsymbol{\epsilon}}^j + \mathbf{E}^j \boldsymbol{\xi})^T \mathbf{\Gamma}^j (\bar{\boldsymbol{\epsilon}}^j + \mathbf{E}^j \boldsymbol{\xi}), \quad (18)$$

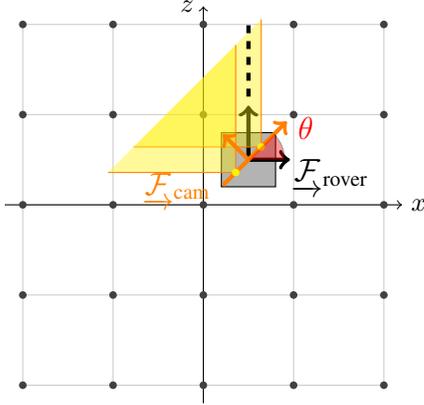


Figure 4: Simulated visual odometry framework. Yellow dots and shaded yellow areas represent the stereo camera and its field of views. Black dots are distributed landmarks and the dashed black line indicates rover motion. The camera is rotated by an angle,  $\theta$ , with respect to the front of the rover. Our paper presents a parametric study of  $\theta$  on the performance of VO.

with

$$\begin{aligned} \bar{\mathbf{e}}^j &= \mathbf{p}_b^j - \mathbf{C}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba}), \\ \mathbf{E}^j &= \left[ \bar{\mathbf{C}}^{ba} - (\bar{\mathbf{C}}_{ba}(\mathbf{p}_a^j - \mathbf{r}_a^{ba}))^\times \right]. \end{aligned}$$

Note that here  $\mathbf{\Gamma}^j$  is defined with  $\bar{\mathbf{C}}_{ba}$  in place of  $\mathbf{C}_{ba}$ .

Taking the derivative of (18) with respect to  $\boldsymbol{\xi}$  gives the following set of linear equations for an update step,  $\boldsymbol{\xi}^*$ :

$$\sum_{j=1}^N \left( \mathbf{E}^{jT} \mathbf{\Gamma}^j \mathbf{E}^j \right) \boldsymbol{\xi}^* = - \sum_{j=1}^N \mathbf{E}^{jT} \mathbf{\Gamma}^j \bar{\mathbf{e}}^j. \quad (19)$$

Once  $\boldsymbol{\xi}^*$  is determined, the state estimate is updated using:

$$\bar{\mathbf{C}}^{ba} \leftarrow e^{-\boldsymbol{\phi}^* \times} \bar{\mathbf{C}}^{ba}, \quad (20)$$

$$\bar{\mathbf{r}}_a^{ba} \leftarrow \bar{\mathbf{r}}_a^{ba} + \boldsymbol{\epsilon}^*. \quad (21)$$

The state is iteratively updated until convergence, which is defined to occur when  $|\boldsymbol{\xi}^*| < 10^{-4}$ .

#### IV. RESULTS

Our simulation involved a rover moving with no wheel slip through a two dimensional world presented in Figure 4, calculating its own egomotion using the VO pipeline discussed in Section III. Zero-mean Gaussian noise was added to the pixel coordinates of all projected landmarks to simulate finite camera resolution. The rover's stereo camera was turned (incrementally) between 0 and 180 degrees relative to the direction of motion.

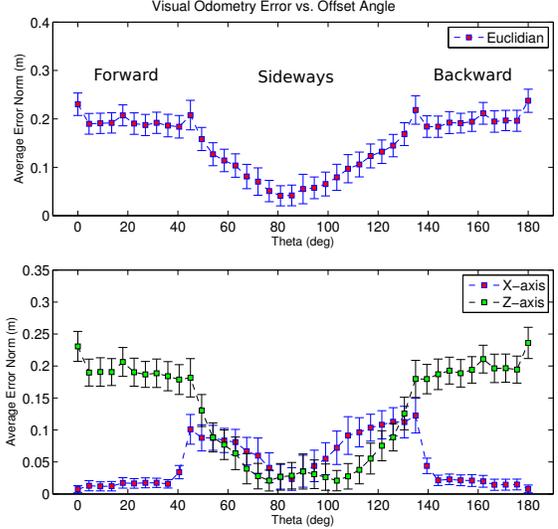


Figure 5: Mean VO estimation error of 500 paths at 40 different camera offset angles ranging from 0 to 180°. The field of view for both camera units was set to 90°, the step size was 1 m and the noise variance was set to 2.25 pixels<sup>2</sup>. The mean error was reduced by 82% when the camera was turned between 85 and 90° to one side. Owing to symmetry, only the range 0° <  $\theta$  < 180° is shown.

#### A. Simulation Results

To begin, we investigated a rover moving in a straight line through a world that contained 400 landmarks distributed on a Cartesian grid at 1 m intervals. VO calculations were performed for two poses separated by 1 m. We repeated the simulation 500 times for each discrete offset angle,  $\theta$ . Results consisting of the mean and standard deviation of the Euclidian position error, as well as the individual components of the error vector, are shown in Figure 5. At an offset angle of 81°, the mean Euclidian error decreased by 82.2% relative to the forward orientation.

We studied a number of factors to determine their effects on this trend. Here, we present two particularly salient ones: (i) the distance between image captures and (ii) camera resolution. The variance of the Gaussian noise added to the pixel coordinates acts as a proxy for camera resolution. Figures 6 and 7 show that the observed trend was most prominent for larger step sizes and lower resolutions.

Further, we investigated the effects of changing the landmark distribution from a Cartesian grid to a randomly distributed set. Figure 8 presents the mean Euclidian error as a function of offset angle,  $\theta$ , for a randomly distributed landmark set. The error follows a similar trend as before, though the symmetry about 90° is often lost. The effect of changing the field-of-view is also illustrated, with larger field-of-views leading to more significant error reductions in transverse camera orientations.

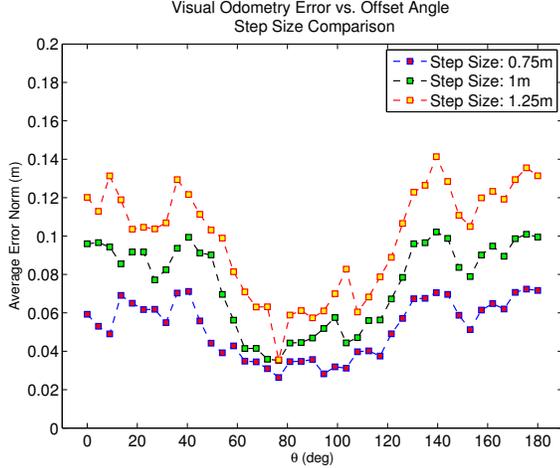


Figure 6: The effects of step size on simulated visual odometry error estimates. Average Euclidian error is plotted. Larger step sizes result in worse VO estimates but greater error reduction at transverse angles.

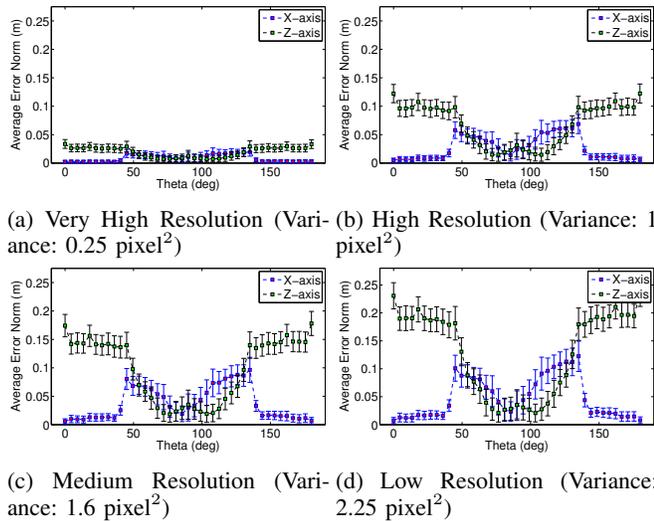


Figure 7: Simulated VO error for four different simulated camera resolutions. Step size is set to 1 m and the field of view to  $90^\circ$ . Increasing the variance (i.e., reducing the resolution) increased the benefit of rotating the camera.

### B. Experimental Results

To validate the simulations, we conducted experimental trials using a Pioneer 3-AT rover outfitted with a Point Grey Bumblebee XB3 stereo camera mounted on a 3-way pan-tilt tripod head. Mimicking the simulations, the camera was oriented parallel to the ground and only its angle about the vertical axis was changed. To establish ground truth, a VICON motion capture system detected a constellation of reflective markers attached directly to the stereo camera. Debris was distributed throughout an indoor facility to

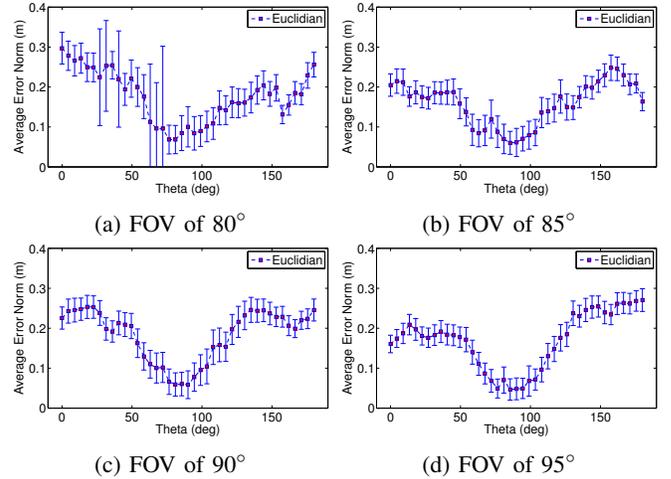


Figure 8: Simulated VO error for a random landmark distribution and four different field-of-view (FOV) values. Step size is set to 1.25 m and noise variance to  $2.25 \text{ pixels}^2$ . A FOV of  $80^\circ$  is required to see significant benefits from rotating the camera.

add supplementary features to the observed environment. Figure 9 illustrates the entire experimental design. In the experiments, we drove the rover 10 meters forward while recording its motion using the VICON system. Figure 10 shows the observed feature tracks for forward and sideways facing camera orientations.

The visual odometry pipeline used Speeded Up Robust Features (SURF) [17] for feature detection and description. Stereo matching was performed by matching SURF descriptors and aided by epipolar constraints. Finally the motion solution was solved using a bundle adjustment technique similar to the matrix weighted approach described in Section IIIB, with a RANSAC procedure used to eliminate outliers.

Figure 11 and Table I show the results of the experiments. Figure 12 shows the position estimates recovered from VO for five different camera orientations. In our data, the mean Euclidian error decreased by a maximum of 59.2% when directing the camera  $135^\circ$  away from the forward direction. The most significant reductions occurred in the regions  $30^\circ < \theta \leq 60^\circ$  and  $105^\circ < \theta \leq 150^\circ$ . The overall error trends were symmetrical about  $\theta = 90^\circ$ , though the backwards-facing orientation (i.e.  $\theta = 180^\circ$ ) exhibited substantially higher errors than directing the camera forwards. VO was performed at 5 Hz, driving at an average speed of 0.5 m/s. The speed resulted in an equivalent step size of around 10 cm. The resolution of each camera within the pair was set to  $512 \times 384$  pixels.

### V. DISCUSSION

The experimental data showed error reductions similar to those from simulations. The optimal angle in the experimental data differed from the simulated optimal ( $135^\circ$  in

Table I: Visual odometry error in experiments. Five trials are averaged at each discrete offset angle.

| Offset Angle ( $^{\circ}$ ) | Mean Path Length | Mean Lateral Error (m) | Mean Forward Error (m) | Mean Euclidian Error (m) | % Improvement |
|-----------------------------|------------------|------------------------|------------------------|--------------------------|---------------|
| 0                           | 9.74             | 0.070                  | 0.11                   | 0.14                     | 0             |
| 45                          | 9.92             | 0.065                  | 0.021                  | 0.071                    | 50.7          |
| 90                          | 9.79             | 0.11                   | 0.087                  | 0.15                     | -4.52         |
| 135                         | 9.86             | 0.049                  | 0.027                  | 0.058                    | 59.2          |
| 180                         | 9.86             | 0.14                   | 0.13                   | 0.21                     | -47.6         |

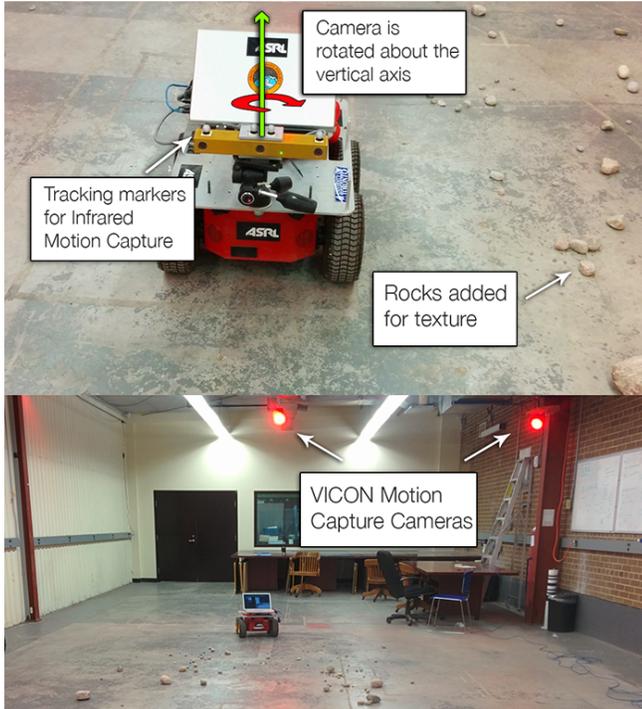


Figure 9: The experimental setup: Pioneer 3-AT rover with a Bumblebee XB3 stereo camera mounted on a pan-tilt head.

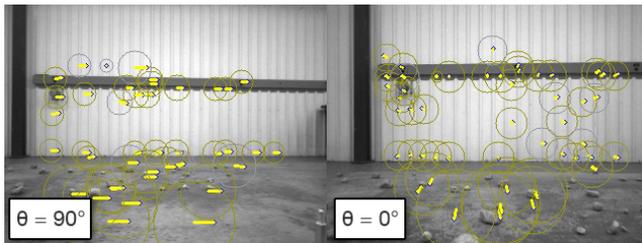


Figure 10: Feature tracks as seen from two different camera orientations ( $\theta = 90^{\circ}$  represents a camera facing to the left of rover motion and  $\theta = 0^{\circ}$  is facing directly ahead).

experimental data, compared to  $82^{\circ}$  in simulations), though as in the simulations, the region  $30^{\circ} < \theta \leq 150^{\circ}$  exhibited significant error reduction. The notable exception is the perpendicular orientation,  $\theta = 90^{\circ}$ . With the camera in this orientation, the mean position error was similar to the front-facing standard. In practice, an oblique offset angle was

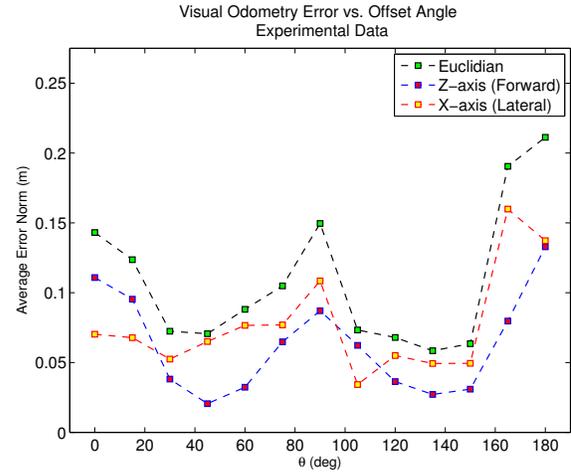


Figure 11: Mean experimental visual odometry error averaged over 5 trials at different angular offsets.

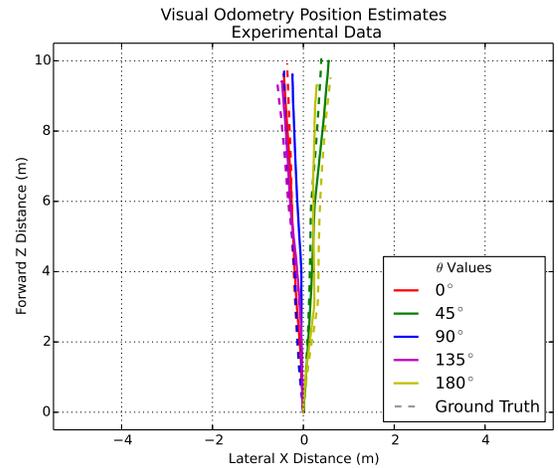


Figure 12: Characteristic experimental visual odometry position estimates for five different offset intervals. Overall mean Euclidian error is significantly reduced at  $\theta = 45^{\circ}$  and  $\theta = 135^{\circ}$ . Oblique angles show growth in lateral error with an accompanied reduction in forward error.

required to minimize the error in position estimates.

The trends in the forward and lateral directions also generally agreed with simulations. As the camera rotated towards

the sideways orientation, lateral error increased slightly or stayed at comparable levels to a forward-facing orientation. Forward error decreased significantly for  $30^\circ < \theta \leq 150^\circ$  with a local maximum at  $90^\circ$ .

A number of different factors could have contributed to the discrepancies between the simulated and experimental data. These include the 3D geometry of our laboratory, the fact that we only report translational differences, and the numerous parameters that were difficult to set arbitrarily and differed from those set in simulations (e.g., field of view, step size, camera resolution, number of tracked landmarks).

Within both the simulations and our experience in the lab, the most significant improvements to VO estimates were observed when computational complexity was limited. The greatest error reduction occurred with a reduced resolution and slower frame rates. We hope that platforms with limited computational resources will benefit from this insight.

## VI. CONCLUSION

In this paper, we showed the effect of changing the orientation of a stereo camera on visual odometry estimates. By rotating the camera about the vertical axis, we reduced visual odometry motion estimation error by as much as 82% in simulations and 59% in real world tests.

Simulations showed that the error was minimized when the camera was rotated approximately 80 degrees relative to the forward direction. In experimental data, we found the most significant error reductions at 45 and 135 degrees. Relatively large step sizes and low resolutions had the largest impact on the trends.

Applications that implement stereo visual odometry with a constrained computational budget may benefit from directing the camera in such a manner. We stress, however, that the trends depend strongly on a number of parameters that are difficult to investigate exhaustively. Nevertheless, in appropriate circumstances, significant error reductions may be achieved with a relatively simple configuration change.

In future work, we plan to explore other state dependent variables that affect optimal stereo camera configuration. This will, in principle, pave the way for an actuated stereo camera that can minimize expected VO error by adapting its orientation based on the rover state and the observed environment.

## ACKNOWLEDGMENT

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) through the NSERC Canadian Field Robotics Network (NCFRN).

## REFERENCES

- [1] D. Nister, O. Naroditsky, and J. Bergen, "Visual odometry for ground vehicle applications," *Journal of Field Robotics*, vol. 23, p. 2006, 2006.
- [2] M. Maimone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the mars exploration rovers," *Journal of Field Robotics*, vol. 24, no. 3, pp. 169–186, 2007.
- [3] 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.
- [4] H. Moravec, "Obstacle avoidance and navigation in the real world by a seeing robot rover," Ph.D. dissertation, Stanford, 1980.
- [5] 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.
- [6] L. Matthies and S. A. Shafer, "Error modeling in stereo navigation," *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 3, 1987.
- [7] F. Fraundorfer and D. Scaramuzza, "Visual odometry : Part i: The first 30 years and fundamentals," *Robotics & Automation Magazine, IEEE*, vol. 18, pp. 80–92, 2011.
- [8] P. T. Furgale, "Extensions to the visual odometry pipeline for the exploration of planetary surfaces," Ph.D. dissertation, University of Toronto, 2011.
- [9] N. Sunderhauf and P. Protzel, "Stereo odometry—a review of approaches," Chemnitz University of Technology, Tech. Rep., 2007.
- [10] B. K. P. Horn, H. M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *J. Opt. Soc. Am. A*, vol. 5, no. 7, pp. 1127–1135, 1988.
- [11] K. Konolige, M. Agrawal, and J. Sola, "Large-scale visual odometry for rough terrain," in *Robotics Research*. Springer, 2011, pp. 201–212.
- [12] 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.
- [13] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376–380, 1991.
- [14] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [15] L. Matthies, "Dynamic stereo vision," Ph.D. dissertation, Carnegie Mellon University, 1989.
- [16] T. D. Barfoot, J. R. Forbes, and P. T. Furgale, "Pose estimation using linearized rotations and quaternion algebra," *Acta Astronautica*, vol. 68, 2011.
- [17] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision—ECCV 2006*. Springer, 2006, pp. 404–417.