

# Estimating Ego-Motion in Panoramic Image Sequences with Inertial Measurements

Felix Schill, Robert Mahony, and Peter Corke

**Abstract** This paper considers the problem of tracking the focus of expansion of a panoramic image sequence due to ego motion of the camera. The focus of expansion provides a measurement of the direction of motion of the vehicle that is a key requirement for implementing obstacle avoidance algorithms. We propose a two stage approach to this problem. Firstly, external angular rotation measurements provided by an on-board inertial measurement unit are used to de-rotate the observed optic flow field. Then a robust statistical method is applied to provide a rough estimate of the focus of expansion as well as a selection of inlier data points associated with the hypothesis. This is followed by a least squares minimisation, utilising only the inlier data, that provides accurate estimates of angular rotation and focus of expansion of the flow. For the robust estimator we consider and compare RANSAC or  $k$ -means segmentation algorithms. The least squares optimisation is solved using a geometric Newton algorithm. The approach is demonstrated on real data obtained from an aerial robotic equipped with panoramic cameras.

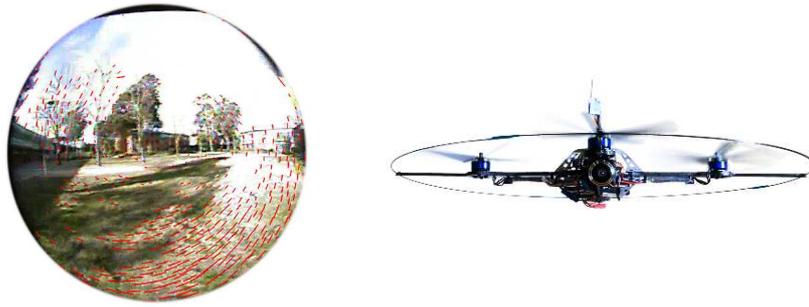
## 1 Introduction

The estimation of the ego-motion of a camera from observation of a sequence of images is a classical problem in the computer vision literature. The classical approach for a projective camera involves recovery of the fundamental matrix between each two images and reconstruction of the motion primitive from these correspondences. Algorithms based on eight or more point correspondences [17, 9] are well known, while for calibrated cameras algorithms exist for as few as five point correspon-

---

Felix Schill and Robert Mahony  
Department of Engineering, Australian National University, ACT, 0200, AUSTRALIA,  
{firstname}.{lastname}@anu.edu.au

Peter Corke  
CSIRO ICT Centre, Pullenvale, Queensland, AUSTRALIA. peter.corke@csiro.au  
**Acknowledgements:** This research was supported by the Australian Research Council through discovery grant DP0880509.



**Fig. 1** Video image and optic flow (left) extracted from on-board video camera of the *Hummingbird* quad-rotor aerial robot (right)

dences [14, 19]. In the case where high accuracy is required the bundle adjustment method can be used [22]. In addition to the classical methods, there have been a wide range of other methods considered in the literature [13, 20, 11, 10, 3, 18].

It is well known that for an image sequence with a small field-of-view it is difficult to distinguish between translation and rotation around orthogonal axes [2, 6]. In addition, there is often a natural bias in solving the instantaneous epipolar constraint, the most common approach to recovering instantaneous motion, associated with grouping of image points in one small area [6]. Using panoramic or catadioptric cameras with a large field-of-view can substantially overcome this issue [6]. Due to the inherent ambiguity in velocity there have been a number of studies based on qualitative analysis of ego-motion methods [4, 5, 21] that utilise panoramic cameras, however, these methods often use explicit search routines to determine the best motion fit and are computationally expensive. In recent work Lim and Barnes have developed methods to compute ego-motion from antipodal pairs of optic flow vectors [15, 16]. Almost all the literature in this area has been developed based on the assumption that the camera is the only sensor. In robotic applications, especially those involving aerial robots, there is almost always an inertial measurement unit (IMU) embarked on the vehicle that can provide a substantially correct estimate of rotation over short periods. However, the vision system for such applications often has poor quality optics, and if the video signal is being transmitted to ground there are artifacts due to signal interference. The authors know of no prior work that addresses the specific issues associated with ego-motion extraction for such a situation.

In this paper, we propose an algorithm for extracting ego-motion from a panoramic image sequence where the angular velocity of rotation can be roughly measured using a separate sensor. We are primarily motivated by applications in aerial robotics where the vehicle is equipped with a wide angle fish-eye (or catadioptric) lens, and inertial sensors. The vision sequences obtained from such vehicles often contain large regions where there is insufficient texture to generate optic flow, for example, regions of sky, or regions distant from the camera where the optics are of insufficient quality to generate good texture. In addition, there are often extreme outliers in the optic flow field caused by errors in the optic flow algorithm induced

by artifacts in the video due to signal interference and multi-path effects in video transmission. We propose a two stage approach. Firstly, optic flow is computed from the image sequence and then this flow is roughly de-rotated using the data from the gyrometer. This is achieved by subtracting the expected rotational optic flow (due to the measured angular velocity) from the measured optical flow. The resulting flow is almost entirely due to the translational ego-motion of the camera, except for errors and noise, and has a simple structure that allows us to develop simple models to determine the unique Focus-of-Expansion (FoE), corresponding to the direction of motion of the camera. We investigate two robust statical methods, RANSAC and K-means, aimed at generating a reasonable hypothesis of the FoE of the flow and identify inlier and outlier optic flow measurements in the data. The K-means algorithm has the advantage that it can potentially identify and segment secondary motion primitives that may be associated with motion of other objects within the scene. We then describe how an initial estimate can be refined in both translation and residual rotation by minimising a least squares cost based on the instantaneous epipolar condition posed on the sphere. We compute the geometric gradient and geometric Hessian and propose a Newton update step. The Newton minimisation is embedded in the RANSAC framework as the second model refinement step for each iteration. Given that the initial estimate provided by the first stage of the algorithm is moderately correct, this stage usually converges in at most three iterations. Moreover, the eigenvalues of the Hessian provide a measure of confidence in the estimate. A poor condition number for the Hessian indicates the likelihood of an unreliable estimate and the overall magnitude of eigenvalues is proportional to distance scaled velocity of the vehicle. The proposed algorithms are tested on synthetic data with outliers and noise, and demonstrated on video and inertial data obtained on a small aerial robotic vehicle.

## 2 Problem formulation

In this section, we introduce some notation and develop the cost function that will be used to refine ego-motion estimates on the sphere.

We are interested in applications where there is a wide field of view fish eye or catadioptric video camera moving through a static world environment. We assume that the camera frame rate is fast compared to the relative optical velocity of the observed scene. Consequently the optic flow can be computed directly on the raw image sequence and the resulting flow vectors mapped back onto a spherical image plane based on a known calibration of the camera. The spherical optical flow field is denoted  $\Phi$  and associates a flow vector  $\Phi(\eta) \in T_\eta S^2$  in the tangent space of the sphere to a point  $\eta \in S^2$  on the sphere. In practice, we are normally constrained to a sparse computation of optic flow, that is measurements at a finite number of points on the sphere that we will index by  $\{\eta_i\}$  for  $i = 1, \dots, n$ , with  $n$  the number of optic flow vectors measured in a given reference image.

The optic flow can be split into a translational and a rotational part

$$\Phi(\eta) = \Psi(\eta) + \Theta(\eta). \quad (1)$$

Here  $\Theta(\eta) := -\Omega \times \eta$ , with  $\Omega \in \{B\}$  the body-fixed frame angular velocity of the camera, is the contribution to the optic flow from the rotational motion of the camera, while the translation component of flow is given by  $\Psi(\eta) := \frac{1}{\lambda(\eta)} \mathbb{P}_\eta v$ , where  $\mathbb{P}_\eta = (I_3 - \eta \eta^\top)$  is the projector onto  $T_\eta S^2$  and  $v \in \{B\}$  is the body fixed frame translational velocity of the vehicle.

In order to derive an optimisation based method for identification of ego motion it is necessary to define a cost function. We propose to use a modified version of the instantaneous epipolar constraint. Let  $\hat{w}$  denote the estimate of  $w \in S^2$  of the true direction of motion of the vehicle. That is, set

$$w = \frac{v}{|v|}, \quad \text{for } v \neq 0,$$

and  $\hat{w} \in S^2$  an estimate of  $w \in S^2$ . The direction of motion  $w$  is also the Focus-of-Expansion (FoE) of the translational flow field  $\Phi$  on the sphere.

For each individual optic flow vector  $\Phi(\eta)$  measured at a point  $\eta \in S^2$  the instantaneous epipolar constraint computed for estimates  $\hat{w}$  and  $\hat{\Omega}$  is

$$e_{\Phi(\eta)}(\hat{w}, \hat{\Omega}) := \langle \hat{w}, (\Phi(\eta) + \hat{\Omega} \times \eta) \times \eta \rangle. \quad (2)$$

Note that if  $\hat{\Omega}$  is correct then  $\Phi(\eta) + \hat{\Omega} \times \eta = \Psi(\eta)$  is the true translational optic flow. Taking the vector product of this with its base point  $\eta$  leads to a vector that is orthogonal to the direction of motion of the vehicle. Taking the inner product of the resulting vector with  $\hat{w}$  is zero precisely when  $\hat{w} = w$  is the true direction of motion. The instantaneous epipolar constraint is often written  $\langle \hat{w} \times \eta, (\Phi(\eta) + \hat{\Omega} \times \eta) \rangle$ , however, this can be transformed into (the negative of) Equation (2) using the properties of vector triple products and the form given above is more convenient for the gradient and Hessian computations undertaken later.

Since the optic flow is measured at a finite number of scattered points the cost considered is a sum

$$f(\hat{w}, \hat{\Omega}) := \sum_{i=1}^n e_{\Phi}^2(\eta)(\hat{w}, \hat{\Omega}) = \sum_{i=1}^n \langle \hat{w}, (\Phi(\eta) + \hat{\Omega} \times \eta) \times \eta \rangle^2 \quad (3)$$

It is clear that for ideal data the cost  $f$  is zero for the correct for  $\hat{w} = w$  and  $\hat{\Omega} = \Omega$ . The cost  $f$  is a smooth function  $f : S^2 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  and can be optimised on this set using geometric concepts. A weakness of the cost proposed is that it is highly susceptible to perturbation by large magnitude outliers in the data. Optic flow algorithms often yield exactly this sort of error due to occasional mismatched point correspondences. Thus, direct minimisation of the cost  $f$  is likely to lead to poor ego-motion estimation. The following section applies robust statistical algorithms to overcome this issue.

### 3 Robust estimation of Focus of Expansion

In this section we present two robust statistical methods for providing an estimate of focus-of-expansion of the image sequence. The approach is directly based on the application domain and we assume that a measurement of angular velocity of the vehicle is available. This is obtained through the inertial measurement unit that is mounted on the aerial vehicles that we consider.

Using the measured angular velocity,  $\Omega_y \in \{B\}$ , the measured optic flow can be de-rotated. That is we define

$$\Psi_{\Omega_y}(\eta_i) := \Phi(\eta_i) - \Theta_{\Omega_y}(\eta_i) = \Phi(\eta_i) + \Omega_y \times \eta_i.$$

The resulting estimate,  $\Psi_{\Omega_y}(\eta_i)$ , of translational flow is only defined at measured flow points  $\eta_i$ .

Any two measurements of translational flow  $\Psi_{\Omega_y}(\eta_i)$  and  $\Psi_{\Omega_y}(\eta_j)$  can be used to generate a hypotheses for the focus of expansion of the flow field. Since the flow vectors must lie in a plane containing the flow vector and the base point  $\eta$  then the intersection of these two planes provides an estimate of the focus-of-expansion for the flow field. Thus,

$$\hat{w}(\Omega_y, \eta_i, \eta_j) := \frac{(\Psi_{\Omega_y}(\eta_i) \times \eta_i) \times (\Psi_{\Omega_y}(\eta_j) \times \eta_j)}{|(\Psi_{\Omega_y}(\eta_i) \times \eta_i) \times (\Psi_{\Omega_y}(\eta_j) \times \eta_j)|}. \quad (4)$$

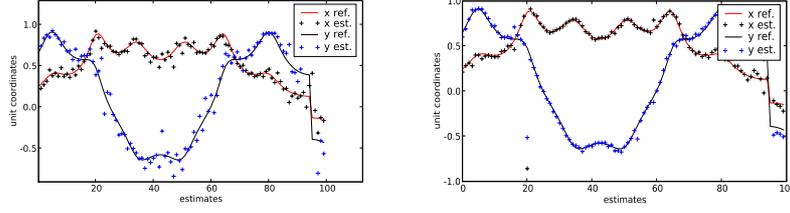
This hypothesis is sign indeterminate so we introduce a corrected hypothesis  $\hat{w}_c$  is

$$\hat{w}_c(\Omega_y, \eta_i, \eta_j) := \hat{w}(\Omega_y, \eta_i, \eta_j) \cdot \text{sign}(\langle \hat{w}, \Phi(\eta_1) \rangle) \quad (5)$$

There are potentially  $n(n-1)/2$  (where  $n$  is the number of flow vectors measured) hypotheses that can be generated based on (4).

The K-means algorithm is based on collecting a large number of hypotheses and then clustering these hypotheses into classes. This is a standard algorithm [8, 12] that is extensively used in the field of computer vision and we will not cover the details of the algorithm implementation. Distance between hypotheses was measured by the cosine of the angle between the two hypotheses. The estimate of the focus-of-expansion is provided by the normalised mean of the largest cluster of hypotheses. Outliers are identified using the clustering identification of the algorithm. The Newton algorithm, described in Section 4 can be applied to the support set of the best cluster from the K-means algorithm to refine the best estimate, however it was found that this would often not converge if there are outliers present (see section 5).

To overcome these limitations, we merged the hypothesis generation of the K-means approach and the Newton model refinement into a more robust RANSAC framework. The RANSAC algorithm is based on consensus scoring of hypotheses and once again is a well known algorithm extensively used in computer vision applications [7, 9]. The algorithm is applied by randomly selecting pairs of flow vectors and generating hypotheses according to Equation 5 and using a normalised mean of the hypotheses generated. All flow vectors are then scored with regard to that hypothesis using cost function 2 to determine the inlier or consensus set, consisting



**Fig. 2** Estimation results for K-means (left plot) and RANSAC (right plot), for synthetic flow, 30% outliers and 0.001 gaussian noise. Note the significantly better performance of the RANSAC algorithm due to the Newton iteration step.

of all flow vectors where  $e_{\Phi(\eta)} < t$  ( $t$  is the acceptance threshold). The hypothesis along with its consensus set is then used to initialise the Newton iteration discussed in Section 4 and a refined estimate of both FoE and angular velocity is computed. Over several iterations, the refined estimate with the smallest residual is chosen as the best estimate from the algorithm.

## 4 Refining the motion estimate

In this section we present a geometric Newton algorithm that can be used to efficiently refine the estimates of ego-motion based on minimising the cost (3). The Newton algorithm requires a reasonable estimate of the local minima and identification of inlier flow vectors to provide a reliable estimate of ego-motion. The implementation of the Newton also needs to respect the unit norm constraint on the focus-of-expansion estimate  $\hat{w}$  in the optimisation problem. We achieve this by deriving the Newton algorithm with respect to the geometry of the constraint set. Details on geometric optimisation algorithms can be found in Absil *et al.* [1].

For the sake of simplifying notation we define

$$Z(\hat{\Omega}) := (\Phi(\eta) + \hat{\Omega} \times \eta) \times \eta. \quad (6)$$

The geometric gradient of  $f$  is an element of  $T_{(\hat{w}, \hat{\Omega})} S^2 \times \mathbb{R}^3$ . It is obtained by differentiating  $f(\hat{w}, \hat{\Omega})$  in an arbitrary direction and then using the natural Riemannian metric to obtain a tangent vector;

$$\text{grad} f(\hat{w}, \hat{\Omega}) = \begin{pmatrix} \mathbb{P}_{\hat{w}} \left( \sum_{i=1}^n Z(\Omega) Z^{\top}(\Omega) \right) \hat{w} \\ - \sum_{i=1}^n \left( (\hat{w}^{\top} Z(\Omega)) \mathbb{P}_{\eta_i} \hat{w} \right) \end{pmatrix}, \quad (7)$$

recalling that  $\mathbb{P}_v = I_3 - vv^{\top}$  is the projection onto  $T_v S^2$ .

It is possible to consider a gradient descent method to optimise the cost function  $f$  (3). However, due to the inherent nature of the data, the cost function is several orders of magnitude more sensitive to change in the angular velocity estimate than the FoE estimate, leading to a highly ill-conditioned optimisation problem. In practice,

effective implementation of a gradient descent algorithm would require preconditioning of the gradient. Since an initial guess of the local minima is available from the K-means or RANSAC algorithm (Section 3) it is possible to overcome this difficulty by using a Newton algorithm directly.

The geometric Hessian for  $f$  can be written

$$\text{Hess}f(\hat{w}, \hat{\Omega}) = \begin{pmatrix} \mathbb{P}_{\hat{w}} \left( \sum_{i=1}^n Z(\Omega) Z(\Omega)^\top \right) \mathbb{P}_{\hat{w}} & -\mathbb{P}_{\hat{w}} \sum_{i=1}^n \left( (\hat{w}^\top Z(\Omega)) \mathbb{P}_{\eta_i} + Z(\Omega) \hat{w}^\top \mathbb{P}_{\eta_i} \right) \\ -\sum_{i=1}^n \left( (\hat{w}^\top Z(\Omega)) \mathbb{P}_{\eta_i} + \mathbb{P}_{\eta_i} \hat{w}^\top Z(\Omega) \right) \mathbb{P}_{\hat{w}} & \sum_{i=1}^n \left( \mathbb{P}_{\eta_i} \hat{w} \hat{w}^\top \mathbb{P}_{\eta_i} \right) \end{pmatrix} \quad (8)$$

The Hessian is written as an element of  $\mathbb{R}^{6 \times 6}$  due to the identification of tangent vectors in  $T_{\hat{w}}S^2$  with elements of  $\mathbb{R}^3$ . However, the vector  $\hat{w}$  is normal to the tangent space  $T_{\hat{w}}S^2$  and it follows that  $v_0 := (\hat{w}, 0) \in \mathbb{R}^6$  is a zero eigenvector of the Hessian  $\text{Hess}f$  in (8). The remaining five eigenvalues are associated with the quadratic structure of the cost  $f$  at the point  $(\hat{w}, \hat{\Omega})$ . Due to the zero eigenvalue the inverse Hessian in the Newton algorithm has to be implemented with a pseudo inverse routine. In addition, the new estimate must be re-normalised onto the sphere at each iteration of the Newton algorithm [1]. For initial conditions close to the minimum of  $f$  each iteration of the Newton algorithm provides an additional two orders of magnitude of accuracy. In practice, at most two or three iterations are sufficient for the purposes of our calculations given that a suitable initial condition is available.

As an additional advantage of applying the Newton algorithm, it is a straightforward exercise to compute the condition number of the Hessian as the ratio of the magnitudes of largest to smallest eigenvalues of the five meaningful eigenvalues of  $\text{Hess}f$  at the cost minimum. The condition number provides a reliability measure for the estimate of ego-motion of the system, a large condition number indicating that the minimisation is highly ill-conditioned. The eigenstructure of the Hessian can be used to identify directions of poor resolution of the ego-motion parameters.

## 5 Results

The combined algorithms were thoroughly tested with synthetically generated optical flow data, and on real video sequences obtained from a small scale quad-rotor aerial vehicle. For the synthetic data, the true ego-motion of the vehicle is known and can serve as ground truth for comparisons. For the video sequences from the flying vehicle the inertial measurements from the on-board IMU were recorded. The measured rotations are used to de-rotate the spherical flow field. The trajectories of the vehicle can be compared qualitatively to the data obtained.

For the simulation tests the flow field was generated by creating a gaussian-distributed point field (offset from origin: 18 along y-axis, sigma=10) for one-sided flow coverage. The offset of the point field simulates incomplete flow from a single camera, covering slightly less than half of the sphere. Tests for surrounding flow coverage were conducted on a gaussian point field centered in the origin with the same variance. Flow is created by translating and rotating the point field, and projecting start- and end position onto the sphere. A certain percentage of vectors

| One-sided flow                |       |       |      |       |      |       |      |       |      |       |
|-------------------------------|-------|-------|------|-------|------|-------|------|-------|------|-------|
| outliers, noise $\rightarrow$ | 0%    | 0.0   | 0%   | 0.001 | 30%  | 0.0   | 30%  | 0.001 | 30%  | 0.002 |
| K-means (translation)         | 0.061 | 0.03  | 2.1  | 1.8   | 1.2  | 1.1   | 3.1  | 2.50  | 5.6  | 4.9   |
| RANSAC (translation)          | 0.001 | 0.001 | 1.0  | 0.9   | 0.5  | 0.003 | 2.2  | 1.6   | 7.7  | 3.4   |
| K-means (with rot.)           | 15.9  | 15.4  | 15.2 | 14.0  | 17.9 | 16.0  | 20.0 | 16.0  | 17.6 | 13.7  |
| RANSAC (with rot.)            | 0.74  | 0.002 | 6.0  | 0.9   | 19.6 | 13.7  | 12.8 | 10.2  | 13.2 | 12.7  |

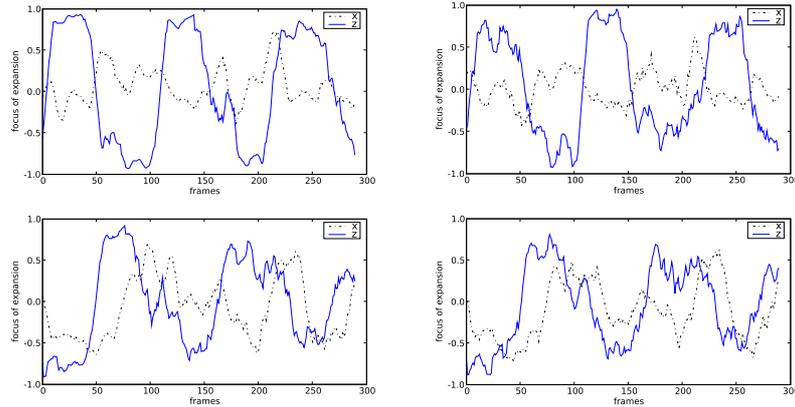
| Surrounding flow              |       |       |      |       |      |       |      |       |      |       |
|-------------------------------|-------|-------|------|-------|------|-------|------|-------|------|-------|
| outliers, noise $\rightarrow$ | 0%    | 0.0   | 0%   | 0.001 | 30%  | 0.0   | 30%  | 0.001 | 30%  | 0.002 |
| K-means (translation)         | 0.03  | 0.015 | 1.50 | 1.0   | 1.4  | 1.1   | 2.5  | 2.0   | 4.0  | 3.0   |
| RANSAC (translation)          | 0.002 | 0.001 | 0.5  | 0.4   | 0.2  | 0.002 | 0.9  | 0.7   | 1.8  | 1.3   |
| K-means (with rot.)           | 13.9  | 11.7  | 14.8 | 10.6  | 17.0 | 11.7  | 13.7 | 10.7  | 12.8 | 11.9  |
| RANSAC (with rot.)            | 3.3   | 0.001 | 1.2  | 0.5   | 12.9 | 6.1   | 11.6 | 4.8   | 14.1 | 7.1   |

**Fig. 3** Mean error (left value) and median error (right value) of FoE in degrees for synthetically generated flow with one-sided flow (upper table) and surrounding flow (lower table) coverage

(30%) is randomised to simulate large outliers, and gaussian noise ( $\sigma = 0.001$  or  $\sigma = 0.002$ ) is added where appropriate in the simulations. The amount of outliers and noise approximately reflect or exceed the distortions found in real optical flow. The results can be seen in figures 3. Parameters for the K-means algorithm were  $k=20$ , and 70 randomly picked pairs. The RANSAC algorithm uses 8 iterations, a threshold  $t = 0.05$ , 20 vector pairs for the initial hypothesis. The Newton algorithm was run for four iterations.

A key observation found was that the performance of the K-means algorithm was not sufficiently reliable to use as the initialisation for the Newton algorithm. In particular, the segmentation of the image flow vectors was not sufficiently robust and the Newton iteration was often undertaken with some outliers that significantly disrupted the performance of the algorithm. As a consequence, the K-means results are presented without any refinement step while the RANSAC algorithm contains the Newton refinement as an integral part. The relative performance of the K-means (without Newton) versus the RANSAC (with Newton) is clearly seen in Figure 2. This can also be seen in the results shown in Table 3. Nevertheless, in the absence of noise all algorithms perform well, even when flow can only be obtained from one side of the sphere. As noise increases, one-sided flow extraction becomes increasingly unstable (notably it is more affected by noise than by outliers). If residual rotation of up to 15 degrees/sec is present in the flow field (i.e. due to imperfect inertial measurements - see lines 3 and 4 in table 3), the estimation results worsen - again more pronounced in the one-sided case.

The algorithms were also applied to video sequences that were collected from a small quad rotor flying vehicle (see Figure 1). The electronically stabilised vehicle is equipped with a forward looking camera with approximately 170 degrees field of view, an IMU, and radio systems that transmit the real-time video images (25 fps) and inertial measurements to the base station. Both signals were synchronised and recorded on the base station. Approximately 100 flow vectors were computed from the image. The test sequences used in this paper consists of the vehicle flying repeatedly forward and backwards by approximately 1.5 m. Two tests were run, the first where the flight was close to linear forward and backward, and the second where the flight was roughly oval. The flight tests were conducted indoors in a confined



**Fig. 4** FoE extraction for real world flyer scenarios. The upper two plots are for linear flight forwards and backwards. The lower two plots are for roughly oval motion of the flyer. The plots on the left show the estimates from K-means, the plots on the right show the RANSAC results.

space, with most obstacles being 1-3 m away from the vehicle. The  $z$ -axis denotes the forward direction of the flyer, the  $x$ -axis points to the right, and the  $y$ -axis points up. Ground truth data was not available for the flight tests. Figure 4 shows the FoE estimates for the flyer with the upper two figures showing the results for linear flight and the lower two showing results for oval flight. The estimates are in vehicle (or camera) coordinates, thus, banking or pitching motions may result in deviations of the estimate from a straight line in world coordinates, however, such motions are small. A sliding window filter of 10 frames was applied to the plots for clarity. The plots show that for real data, both the K-means and RANSAC algorithm deliver reasonably good estimates of the direction of travel. The additional advantages of the RANSAC algorithm are somewhat impacted by the quality of the data.

## 6 Conclusions

Two methods were presented for estimating the focus of expansion from sparse panoramic optical flow fields, namely a K-means clustering method, and a RANSAC framework using a Newton iteration for model fitting. The authors introduced a cost function, gradient and Hessian for estimation of direction of travel, and ego-rotation, which enables gradient descent methods and Newton methods for estimating ego-motion from spherical optic flow. The presented methods work with sparse, patchy flow fields, even if less than half the sphere is covered. Measurements from inertial sensors are used to provide a good initial value in rotation for the algorithms. Removing rotation from the flow field improves the results significantly. The algorithms were evaluated and compared on synthetic data; it was found that the RANSAC algorithm performs better, but also that the K-means algorithm provides good results at much less computational cost. Tests on video and inertial measure-

ments from a quad-rotor flying vehicle show that both algorithms can be applied to real data obtained from a single fish-eye camera, and provide a good estimate of the direction of travel.

## References

1. P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, New Jersey, USA., Princeton, NJ, January 2008. ISBN 978-0-691-13298-3.
2. G. Adiv. Inherent ambiguities in recovering 3-d motion and structure from a noisy flow field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:477–489, 1989.
3. A. Agrawal and R. Chellappa. Ego-motion estimation and 3d model refinement in scenes with varying illumination. In *IEEE Workshop on Motion and Video Computing, (MOTIONS 05)*, volume 2, pages 140–146, 2005.
4. T. Brodsky, C. Fermuller, and Y. Aloimonos. Directions of motion fields are hardly ever ambiguous. *International Journal of Computer Vision*, 26(1):5–24, 1998.
5. C. Fermuller and Y. Aloimonos. Qualitative egomotion. *International Journal of Computer Vision*, 15:7–29, 1995.
6. C. Fermuller and Y. Aloimonos. Ambiguity in structure from motion: Sphere versus plane. *International Journal of Computer Vision*, 28(2):137–154, June 1998.
7. 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*, 24:381–395, 1981.
8. J.A. Hartigan and M.A. Wong. Algorithm as 136: A k-means clustering algorithm. *Applied Statistics*, 28(1):100–108, 1979.
9. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
10. M. Irani, B. Rousso, and S. Peleg. Recovery of ego-motion using region alignment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(3):268–272, 1997.
11. A. Jepson and D. Heeger. Subspace methods for recovering rigid motion i: algorithm and implementation. *International Journal of Computer Vision*, 7(2), 1992.
12. T. Kanungo, D.M. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A.Y. Wu. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:881–892, 2002.
13. J. Koenderink and A. van Doorn. Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22(9):773–791, 1975.
14. H. Li and R. Hartley. Five-point motion estimation made easy. In *18th International Conference on Pattern Recognition (ICPR06)*, pages 630–633, 2006.
15. J. Lim and N. Barnes. Estimation of the epipole using optical flow at antipodal points. In *OMNIVIS*, 2007.
16. John Lim and Nick Barnes. Directions of egomotion from antipodal points. In *Proceedings of CVPR*, 2008.
17. H. Longuet-Higgins. A computer algorithm for reconstruction of a scene from two projections. *Nature*, 293:133–135, 1981.
18. A. Makadia, C. Geyer, S. Sastry, and K. Daniilidis. Radonbased structure from motion without correspondences. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05)*, June 2005.
19. D. Nister. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 756–770, 2004.
20. J.W. Roach and J. K. Aggarwal. Determining the movement of objects from a sequence of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(6):55–62, 1980.
21. C. Silva and J. Santos-Victor. Direct egomotion estimation. In *Proceedings 13th International Conference on Pattern Recognition*, volume 1, pages 702–706, 1996.
22. B. Triggs, P.F. McLauchlan, R.I. Hartley, and A.W. Fitzgibbon. Bundle adjustment - a modern synthesis. In *In Vision Algorithms 99*, pages 298–372, 1999.