

Feature Based Potential Field for Low-level Active Visual Navigation

Rômulo T. Rodrigues¹, Meysam Basiri¹, A. Pedro Aguiar², and Pedro Miraldo¹

¹ Institute for Systems and Robotics, Lisbon, Portugal,
 {romulo.rodrigues, meysam.basiri, pedro.miraldo}@tecnico.ulisboa.pt,

² Faculty of Engineering University of Porto, Porto, Portugal,
 pedro.aguiar@fe.up.pt,

Abstract. This paper proposes a novel solution for improving visual localization in an active fashion. The solution, based on artificial potential field, associates each feature in the current image frame with an attractive or neutral potential energy. The resultant action drives the vehicle towards the goal, while still favouring feature rich areas. Experimental results with a mini quadrotor equipped with a downward looking camera assess the performance of the proposed method.

Keywords: aerial systems, perception and autonomy, autonomous vehicle navigation, localization

1 Introduction

To enable the deployment of autonomous Micro Aerial Vehicles (MAVs) in GPS-denied environments or in scenarios where external systems such as motion tracking cameras are not available, the localization problem, that is, the problem of estimating the pose of the MAV with respect to its environment, needs to be adequately addressed. Research community has been focused in using lightweight onboard sensors. Most effective solutions combine Visual Odometry (VO) algorithms and inertial data within a filtering framework [1,2,3].

Basically, visual odometry aims at estimating the camera motion by taking the motion of tracked features in two consecutive frames into consideration. Key-point methods [4] extract features from salient image regions, to recover camera pose using epipolar geometry. While direct methods [5,6] consider photometric information at high frame rates for a robust pose estimation. For passive VO [7], interest image regions are initialized and tracked as a consequence of navigation. When visual cues are not available, the vision pipeline fails and state estimation relies on model propagation and inertial data. Within few seconds, state estimation error and uncertainty grows fast in time and vehicle may get lost, as is illustrated in Fig. 1(a). There are several approaches in the literature, namely active-SLAM [8,9], next-best-view [10,11] and planning under uncertainty [12,13] that tackle this problem. The basic idea consists in adding visual information in the motion planning and control loop to minimize state estimation uncertainty.

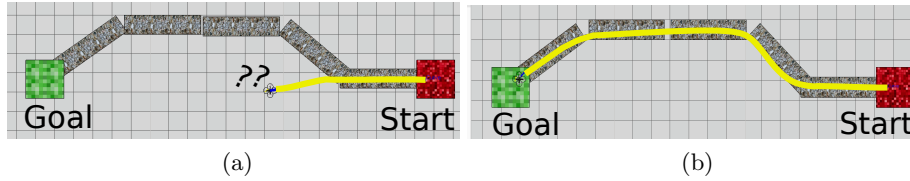


Fig. 1: Localization may fail when not considering the perception pipeline in the motion planning and control frameworks (a). The proposed solution computes an actuation that drives the vehicle into feature rich zones, avoiding localization failure. (b). Illustrations were made using V-REP by Coppelia Robotics.

Davison and Murray [8] define the main goal in active feature selection as building a map of features, which helps localizing the robot rather than an end result in itself. In that sense, Vidal-Calleja et al. [9] propose a control law that drives the camera such that expected information gain is maximized. Sadat et al. [10] propose a scoring function that takes into account the expected number of features for a given camera’s viewpoint, using a mesh of triangles. Mostegel et al. [11] propose a set of measurements, including geometric point quality and recognition probability, to analyze the impact of possible camera motions, and avoid localization loss.

Considering a known map (given *a priori*), the navigation task can be improved by planning routes that favor texture-rich areas. Achtelik et al. [12] addresses a Rapid-exploring Random Belief Tree (RRBT) framework that incorporate MAV dynamics and pose uncertainty. In [13], pose uncertainty is incorporated in a Rapid-exploring Random Trees (RRT*) framework. Most informative trajectories are selected using Fischer informative matrix. Previous non-mapped or non-static regions have impact on local edges and vertexes affected by new information.

This paper proposes a low-level strategy for MAVs to improve the locally visual odometry performance. Therefore, high level localization task, such as Simultaneous Localization And Mapping (SLAM) are also benefited. The solution guides the vehicle through a path with high quality image features. In particular, we propose a method based on Artificial Potential Field (APF) where each image feature is associated with a corresponding potential energy. A reference velocity is thus derived that points towards a feature rich region in the current image frame. The total control input velocity is a linear combination of the reference velocity with the one imposed by the motion control algorithm so that the vehicle is driven toward the goal, while avoiding texture-less and non-static features regions, as it is illustrated in Fig. 1(b).

Most active localization solutions addressed in the literature are built on top of SLAM, being relatively complex and map dependent. In contrast, our method does not require a map, but only features selected as inliers in the current frame. Therefore, similarly to the original work on artificial potential fields by Khatib

[14], it is a low level real-time local solution which can be integrated with a high-level method for better performance. In particular, the proposed solution will not experience the same type of errors as in active SLAM solutions and can be integrated in a complementary fashion such that the final solution avoids local minima.

The remainder of this paper is organized as follows: Sec. II introduces basic notations and definitions. In Sec. III, the proposed method is presented. Sec. IV presents experimental results. Finally, Sec. V addresses final remarks and future work.

2 Notations and definitions

Consider the 3D body fixed frame $\{B\}$ and the 2D image plane frame $\{I\}$. The origin of $\{B\}$ coincides with the center of gravity of the vehicle and the origin of $\{I\}$ corresponds to the top-left image pixel. Vectors are described in lower case bold and a leading superscript indicates its coordinate frame. The homogeneous coordinate of vector \mathbf{v} is denoted as $\bar{\mathbf{v}}$. When a vector is described in $\{I\}$, the leading superscript is omitted. Matrices are written in upper case and sets in calligraphic letter. The transformation from body to image frame $T = K[R|\mathbf{t}] \in \mathbb{R}^{3 \times 4}$ is known, where R and \mathbf{t} denotes the rotation and translation from $\{B\}$ to $\{C\}$, and K is the intrinsic parameter matrix of the camera.

A high level positioning controller computes ${}^B\mathbf{v}_g = (v_x, v_y, 0) \in \mathbb{R}^3$, a velocity that drives the vehicle to the spatial goal. Notice that only planar motion is considered, that is the vehicle keeps a constant height. Also, let $\mathbf{p} = (u, v) \in \mathbb{R}^2$ be the undistorted coordinates of an image feature. In particular, $\mathbf{p}_i \in \mathcal{F}$ is the set of features tracked and selected as inliers in the current image frame.

3 Proposed Method

In this section we describe the proposed method for computing the velocity that drives the vehicle towards a spatial goal, while avoiding low feature areas. This is achieved by adding a component to the goal velocity vector ${}^B\mathbf{v}_g$ that favors rich regions regarding features.

3.1 Features to charge

Each feature is associated with an attractive or neutral potential energy. Associating similar potential energy to every feature in the image frame is not adequate, since the system could be easily trapped in a local minima or subject to rough changes when new features are extracted. Instead, taking advantage of the fact that the camera provides bearing information, the proposed method considers the orientation of each feature w.r.t. the direction of \mathbf{v}_g , i.e. the projection of the goal velocity in the image frame. Thus, as it is expected in a potential field framework, the final goal itself plays a part in the local decision making process.

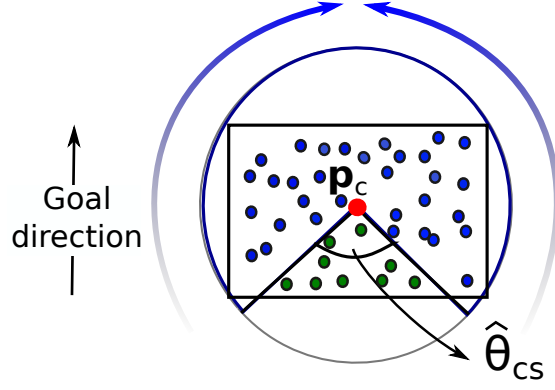


Fig. 2: Attractive potential energy increase in the direction of the blue arrow. Charges are neutral within the circle segment defined by θ_{cs} . Attractive and neutral charges are represented by blue and green dots, respectively.

Let $\mathbf{p}_c = (u_r, v_r)$ be a point that belongs to the image frame, and consider that the feature based velocity shall be computed at that point. For each feature $\mathbf{p}_i \in \mathcal{F}$, compute

$$\hat{\mathbf{p}}_i = (\hat{u}_i, \hat{v}_i) = \mathbf{p}_i - \mathbf{p}_c \quad (1)$$

$$\theta_i = \arccos \left(\frac{\langle \hat{\mathbf{p}}_i, \mathbf{v}_g \rangle}{\|\hat{\mathbf{p}}_i\| \|\mathbf{v}_g\|} \right), \quad (2)$$

where $\theta_i \in [0, \pi]$ and $\mathbf{v}_g \sim T^B \bar{\mathbf{v}}_g$. Notice that only the direction of ${}^B \mathbf{v}_g$ is taken into account for computing the feature-based velocity vector. The last coordinate of its homogeneous form must be set to 0.

Without loss of generality, features in the image plane can be confined within the boundary of a circle centered at \mathbf{p}_c , as shown in Fig. 2. Let the central angle $\hat{\theta}_{cs}$ define a circular segment in the circle, such that the angle θ_{cs} is defined as

$$\theta_{cs} = \pi - \hat{\theta}_{cs}/2. \quad (3)$$

A charge $Q_i \in \mathcal{Q}$ is represented as the tuple $Q_i = (\hat{\mathbf{p}}_i, q_i, \theta_{cs})$, where $q_i \in [0, 1]$ is its corresponding potential energy. The charging policy is defined below:

$$q_i = \begin{cases} 1 - \frac{\theta_i}{\theta_{cs}}, & \text{if } \theta_i \leq \theta_{cs} \\ 0, & \text{if } \theta_i > \theta_{cs} \end{cases}. \quad (4)$$

Based on the current frame information, the charging policy considered associates high attractive potential energy to features localized in the goal direction. Features localized away from the goal direction, on the region defined by the circle segment, have a neutral charge.

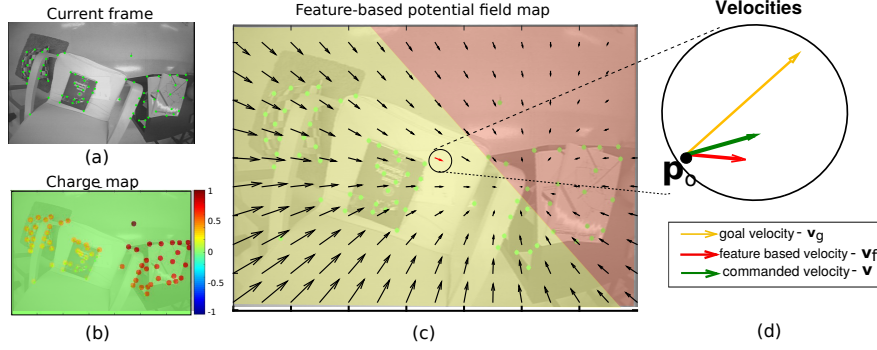


Fig. 3: Case study for the feature based potential field. Current image frame and inlier features (a). Considering \mathbf{p}_c as the optical center, the charge map is built and represented as a heat-map. (b). The potential field map shows the action derived when evaluating at different \mathbf{p}_c (c). Goal-friendly region is shown in yellow and feature-friendly in red (more detail about these regions is shown in Sec. 3.3). (d) shows commanded actuation as the combination of goal and feature velocity. ($\mathbf{p}_c = \mathbf{p}_o$, $\lambda = 0.1$, $\theta_{CS} = 10^\circ$, $d = 50$ pixel, $s = 150$ pixel).

3.2 Vector Field

Each charge $Q_i \in \mathcal{Q}$ exerts a force \mathbf{f}_i at \mathbf{p}_c , given by

$$\mathbf{f}_i = \begin{cases} (0, 0)^T, & \text{if } d_i < r \\ \frac{(d_i - r)}{s} q_i \begin{pmatrix} \cos(\phi_i) \\ \sin(\phi_i) \end{pmatrix}^T, & \text{if } r \leq d_i \leq s + r \\ q_i \begin{pmatrix} \cos(\phi_i) \\ \sin(\phi_i) \end{pmatrix}^T, & \text{if } d_i > s + r \end{cases}, \quad (5)$$

where r is the distance in pixels that a charge must be, from the evaluated point, to exerted any force on it; s is the spread in pixels of the potential field, and d_i and ϕ_i are computed as follow

$$d_i = \|\hat{\mathbf{p}}_i\| \quad (6)$$

$$\phi_i = \text{atan2}(\hat{u}_i, \hat{v}_i). \quad (7)$$

The total force \mathbf{f} on the point \mathbf{p}_c can be computed as

$$\mathbf{f} = \sum_i \mathbf{f}_i, \quad (8)$$

which can be normalized and transformed into a feature based velocity command $\bar{\mathbf{v}}_f$. Its direction in homogeneous coordinates is given as

$$\bar{\mathbf{v}}_f = \frac{1}{\|\mathbf{f}\|} \begin{bmatrix} \mathbf{f} \\ 0 \end{bmatrix}. \quad (9)$$

Finally, the proposed command action takes the form

$$\bar{\mathbf{v}} = \lambda \bar{\mathbf{v}}_g + (1 - \lambda) \bar{\mathbf{v}}_f, \quad (10)$$

where λ is a weight factor and $\bar{\mathbf{v}}_g$ is a normalized velocity. The velocity $\bar{\mathbf{v}}$ can be transformed from the image frame to the body frame and scaled accordingly.

3.3 Discussion

Artificial potential fields frameworks usually take into consideration the robot position when computing forces - obstacles exert repulsive force and the goal an attractive force. In the proposed framework, features can be either attractive or neutral accordingly to their position w.r.t. the point \mathbf{p}_c being considered and the goal direction. As for now, only attractive and neutral charges are admitted, in the future, the effect of associating repulsive charge to features classified as outlier will be analyzed.

Figure 3 illustrates a case study. In particular, Fig. 3(a) shows a frame and extracted features classified as inliers. The goal velocity \mathbf{v}_g is directed towards the top-right pixel in the image frame - a poor zone regarding the number of features. Fig. 3(b) depicts the potential energy associate to each feature when evaluating the action induced at the central pixel of the camera, \mathbf{p}_o . The potential field map (see Fig. 3(c)) shows the corresponding field for different values of \mathbf{p}_c . For each point in the map, (1)–(10) must be computed. However, for visualization, \mathbf{v}_f is not normalized. The map can be classified in a *goal-friendly* and *feature-friendly* actuation zone. Both regions are shown in the potential field map in yellow (goal-friendly) and red (feature-friendly) background. Suppose $\lambda = 0$, then according to (10) the vehicle follows \mathbf{v}_f . If \mathbf{p}_c is within a feature-friendly region, the vehicle favors more the features than the goal. As a matter of fact, the vehicle will move away from the goal. On the contrary, if \mathbf{p}_c is within a goal-friendly region, the vehicle will move towards the goal. The radius of the charge r determines whether features close to the point \mathbf{p}_c affect the solution or not. Meanwhile, the spread s determines the strength of each charge. The larger the spread is, the more influence charges away from the point evaluated will have on the feature driven action. Thus, it limits the prediction horizon based on the local frame, that is, the belief that a feature on an edge indicates there will be more features on that direction.

4 Experimental Results

This section describes the experimental setup, as well as the results observed for multiple flight tests.

4.1 Experimental Setup

Fig. 4 shows the main components employed in experimental validation. The algorithm was tested in the mini-quadrotor Crazyflie 2.0, manufactured by BitCraze.

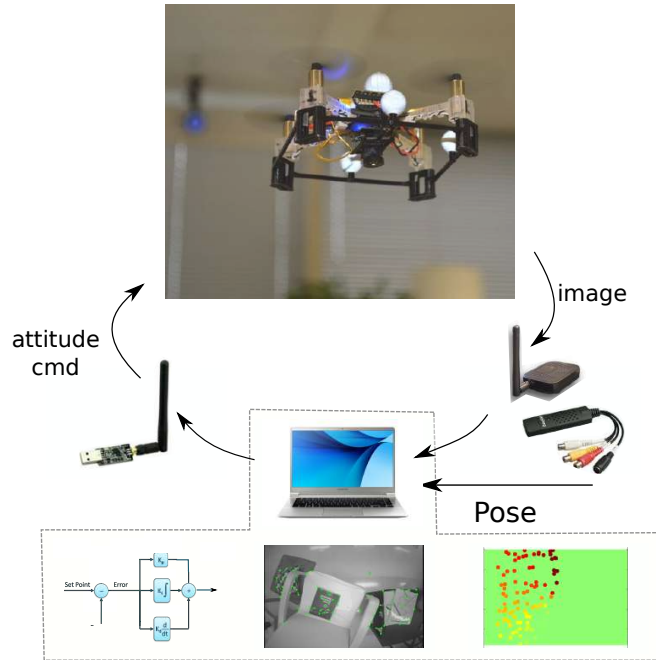


Fig. 4: Experimental setup. Clockwise, starting from the top: the crazyflie equipped with the mini transmitter camera module FX798T, image receiver, and capture card; a notebook that runs the code; and an antenna for the communication with the vehicle.

We attached to the vehicle a 4.7 g mini transmitter camera module FX798T equipped with 120° field of view lens - the camera faces downward. This module broadcasts images using an embedded 5.8 GHz transmitter. On the other side, we use a receiver that is connected to a computer through a video capture card. The frame-rate is 30 Hz for a 720 × 480 image resolution. The proposed algorithms runs on the off-board computer at frame-rate speed. The flying arena is equipped with an Optitrack, motion capture system manufactured by the company NaturalPoint. The system records the motion of retro-reflective markers using the optical-passive technique. Infrared cameras capture markers' position allowing a pose estimation with millimetric precision.

Images and commands are published and received within the ROS (Robot Operating System) environment. Collective thrust and attitude commands are sent to the vehicle using the package developed in [15]. A customized PID controller assures the vehicle follows the desired action. The motion capture system provides feedback for the control loop, while the feature-based potential field works online as the decision making process, sending desired velocity commands to the controller. A first order low pass filter with cutoff frequency at 20 Hz provides smooth control reference for the feature driven action.

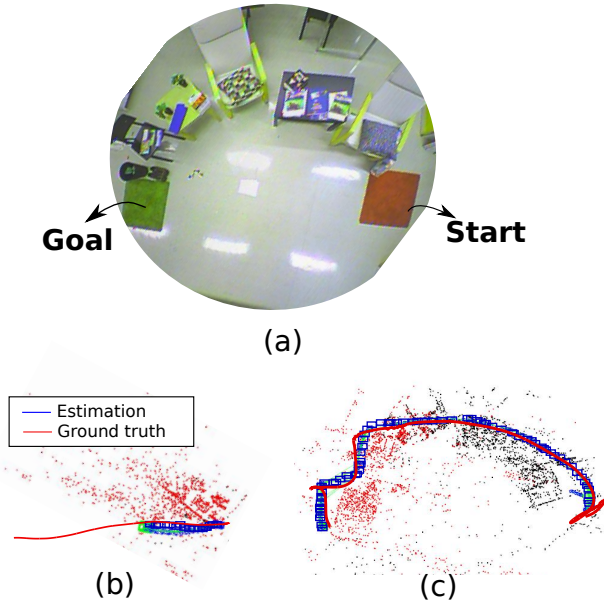


Fig. 5: Scenario for evaluation of proposed algorithm (a). Localization fails in the absence of features (b). Feature based velocity vector drives the vehicle to the goal through a region rich in features (c).

When a new image frame is received, Shi-Tomasi features [16] are extracted. In addition to a minimum quality threshold, only 100 features with the highest response are selected. Then, features are tracked across two consecutive frames using Lucas-Kanade Tracker (LKT) [17]. Within a RANSAC [18] framework, the 8-point algorithm classifies features as inliers or outliers. The solution is robust under the presence of few false inliers, such that a matching step is unnecessary.

4.2 Evaluation

Multiple flights were performed in a $4\text{ m} \times 3\text{ m}$ living room alike arena, as shown in Fig. 5(a). The robot was autonomously controlled to fly from the starting position to the goal position as indicated by the two carpets. As seen from the image, the straight path that links both carpets is a regular floor that does not contain good visual cues. The localization task for a trial is said to *succeed* if ORB-SLAM [4] manages to keep track of the pose of the vehicle. ORB-SLAM run offline, just for evaluation purpose. Multiple trials showed that taking the straight path towards the goal always results in localization failure due to the lack of visual features. Fig. 5(b) shows one of the trials where the localization estimation fails when a straight path is taken. Alternatively, using the proposed active method, the robot was always capable of reaching the goal while maintaining localization. Fig. 5(c) shows the path the robot takes, as well as the

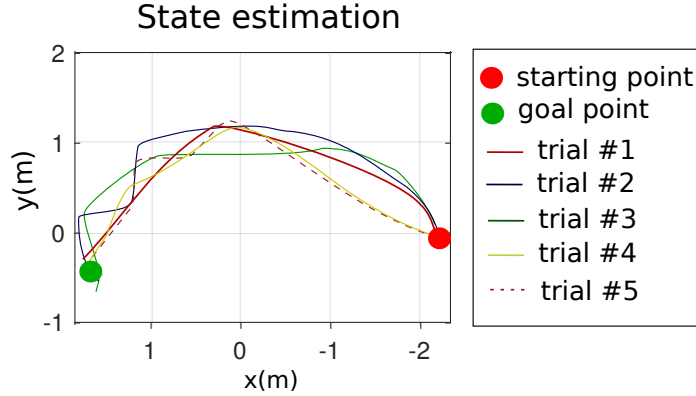


Fig. 6: Multiple trial for $\lambda = 0.45$, $\hat{\theta}_{cs} = 30^\circ$ and $\mathbf{p}_c = \mathbf{p}_o$.

estimated state throughout the time in one of the experiments. Fig. 6 shows different trials with the proposed method where the robot reaches the goal while keeping track of its location.

5 Conclusions

This paper proposes a low-level solution for the active visual navigation problem. Features tracked across consecutive frames are associated with an attractive or neutral potential energy. The intensity of each charge is a function of the goal direction, which is given by a motion-control algorithm. The desired vehicle velocity combines a component that takes visual cues into consideration and a component related to the spatial goal. Experimental results using a micro aerial vehicle, equipped with a downward looking camera, showed that the method can effectively drive the vehicle towards the goal while avoiding no or poor featured regions. The proposed active solution does not rely on a map and hence can be integrated within a SLAM framework to improve the accuracy and robustness of localization and mapping. Points that are currently being investigated include the analysis of the effect of dynamic tuning of parameters based on mission & environment characteristics, associating repulsive energy to features classified as outliers, and the reasoning about computing the feature based velocity vector in other points than the camera's optical center.

Acknowledgement

R. Rorigues, M. Basiri, and P. Miraldo were partially supported by Fundação para a Ciência e Tecnologia (FCT) project UID/EEA/50009/2013, and by FCT grant SFRH/BPD/111495/2015. P. Aguiar was partially supported by project

POCI-01-0145-FEDER-006933/SYSTEC funded by FEDER funds through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI) - and by national funds through FCT.

References

1. Loianno, G., Watterson, M., Kumar, V.: Visual inertial odometry for quadrotors on SE(3). In: IEEE Int'l Conf. on Robotics and Automation (ICRA), pp. 1544-1551 (2016).
2. Weiss, S., Achtelik, M. W., Lynen, S., Achtelik, M. C., Kneip, L., Chli, M., Siegwart, R.: Monocular Vision for Long-term Micro Aerial Vehicle State Estimation: A Compendium. In: Journal of Field Robotics, **30**(5):803-831 (2013).
3. Engel, J., Sturm, J., Cremers, D.: Accurate Figure Flying with a Quadcopter Using Onboard Visual and Inertial Sensing. In: Proc. of the Workshop on Visual Control of Mobile Robots (ViCoMoR) (in conjunction with IROS'12) (2012).
4. Mur-Artal, R., Montiel, J. M. M., Tardes, J. D.: ORB-SLAM: A Versatile and Accurate Monocular SLAM System. In: IEEE Trans. on Robotics, **31**(5):1147-1163 (2015).
5. Forster, C., Pizzoli, M., Scaramuzza, D.: SVO: Fast Semi-Direct Monocular Visual Odometry. In: IEEE Int'l Conf. on Robotics and Automation (ICRA), pp. 15-22 (2015).
6. Engel, J., Koltun, V., Cremers, D.: Direct Sparse Odometry. In: arXiv:1607.02565 [cs.CV] (2016).
7. Scaramuzza, D., Fraundorfer, F.: Visual Odometry [Tutorial]. In: IEEE Robotics Automation Magazine, **18**(4):80-92 (2011).
8. Davison, A. J., Murray, D. W.: Simultaneous localization and map-building using active vision. In: IEEE Trans. on Pattern Analysis and Machine Intelligence, **24**(7):865-880 (2002).
9. Vidal-Calleja, T. A., Sanfeliu, A., Andrade-Cetto, J.: Action selection for single-camera SLAM. In: IEEE Trans. on Systems Man and Cybernetics, Part B (Cybernetics), **40**(6):1567-1581 (2010).
10. Sadat, S. A., Chutskoff, K., Jungic, D., Wawerla, J., Vaughan, R.: Feature-rich path planning for robust navigation of MAVs with Mono-SLAM. In: IEEE Int'l Conf. on Robotics and Automation (ICRA), pp. 3870-3875 (2014).
11. Mostegel, C., Wendel, A., Bischof, H.: Active monocular localization: Towards autonomous monocular exploration for multirotor MAVs. IEEE Int'l Conf. on Robotics and Automation (ICRA), pp. 3848-3855 (2014).
12. Achtelik, M. W., Lynen, S., Weiss, S., Chli, M., Siegwart, R.: Motion- and Uncertainty-aware Path Planning for Micro Aerial Vehicles. In: Journal of Field Robotics, **31**(4):676-698 (2014).
13. Costante, G., Forster, C., Delmerico, J., Valigi, P., Scaramuzza, D.: Perception-aware Path Planning. In: arXiv:1605.04151 [cs.RO] (conditionally accepted for IEEE Trans. on Robotics) (2016).
14. Khatib, O.: The Potential Field Approach And Operational Space Formulation In Robot Control. In: Adaptive and Learning Systems: Theory and Applications, 367-377 (1986).
15. Hoenig, W., Milanes, C., Scaria, L., Phan, T., Bolas, M., Ayanian, N.: Mixed Reality for Robotics. In: /RSJ Int'l Conf. Intelligent Robots and Systems. pp. 5382-5387 (2015).

16. Shi, J., Tomasi, C.: Good features to track. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 593-600 (1994).
17. Lucas, B. D., Kanade, T.: An Iterative Image Registration Technique with an Application to Stereo Vision. pp. 674-679 (1981).
18. Fischler, M. A., Bolles, R. C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. In: Commun. ACM. **24**(6):pp. 381-395 (1981).