An Information-Theoretic Active Localization Approach during Relative Circumnavigation in Orbit

Michail Kontitsis*, Panagiotis Tsiotras†, and Evangelos A. Theodorou‡

Georgia Institute of Technology, Atlanta, GA, 30332, USA

This paper presents an information-theoretic active localization technique applied to the problem of relative navigation and self-localization in orbit. We apply the approach to the problem of a chaser satellite circumnavigating a target satellite in a circular orbit, while observing a set of feature points (landmarks) on the target satellite. The approach relies on the Cross Entropy (CE) optimization method to select camera orientation trajectories that minimize both the localization error and the corresponding uncertainty bounds, while moderating the required control effort. The proposed method provides a framework for near-optimal solutions by jointly considering control, planning and estimation. We show the benefits of the method in terms of landmark uncertainty reduction and we compare the proposed approach against an open-loop strategy and a "greedy" feedback strategy.

I. Introduction

Autonomous satellite relative navigation has received a lot of attention lately owing to its potential impact on proximity operations in orbit. In fact, autonomous proximity operations (e.g., monitoring, servicing, docking) are identified in NASA's roadmap[1] as an area of great interest for many future missions, both in low Earth orbit (LEO) as well as in deep space. Fundamental to the achievement of such autonomy is the ability to accurate localize a spacecraft with respect to another spacecraft or asteroid. Existing work in proximity operations (see, for example Refs. [2–6]) solve the problem of control and estimation independently. That is, various landmarks, such as patterns or features on the target satellite or asteroid, are first localized and are subsequently used as reference points for other control tasks, such as docking or precision landing. The effect that the particular control policies might have on the estimation and vice-versa, by and large, are not explored. Instead, the standard approach is based on the separation principle of stochastic optimal control,[7–9] according to which the control and estimation problems are dual and they can be solved completely separately from one another. This, however, is true only for linear systems with additive Gaussian process and observation noise.

In this work we integrate planning and stochastic optimization with agent localization in order to perform control under uncertainly for an autonomous spacecraft in orbit. Control here means the active rotation of the spacecraft or sensor about one of its axes so that it points to the location of the feature point(s) detected in a previous time step. The estimation accuracy of the detected feature(s) drives which feature(s) should be selected next, and hence also drives the corresponding control action. The control and estimation steps are therefore coupled. This active localization approach, when coupled with a mapping step of the initially unknown environment, is known in the robotics community as Active SLAM (Simultaneous Localization and Mapping) and it has been studied extensively in the past, mainly in the robotics community.[10–12] In particular, in Refs. [11, 12] model predictive control has been used for planning trajectories which improve SLAM performance. Simulations and experiments were performed for a single agent exploration scenario. In Ref. [10] Active SLAM was performed for the case of single robot while the metric used for planning was the information gain. For some more recent results on the subject, see Refs. [13–15].

In this paper, and in order to capture estimation uncertainty in the optimization step, we construct a cost function which includes, besides the state and control cost, a term encapsulating the uncertainty of the state

*Postdoctoral Fellow, School of Aerospace Engineering.

†Dean's Professor, School of Aerospace Engineering, and Institute for Robotics and Intelligent Machines, AIAA Fellow.

‡Assistant Professor, School of Aerospace Engineering, and Institute for Robotics and Intelligent Machines.

as provided by an estimator, e.g., an Extended Kalman Filter (EKF) or Unscented Kalman Filter (UKF). The resulting optimization problem is difficult to solve. We thus resort to the method of cross-entropy (CE) minimization to find the optimal control strategy. The CE minimization method iteratively selects the best attitude trajectories to minimize the aforementioned cost. The result is a near-optimal path, in terms of achieving a predefined goal in the state space, while reducing the localization error and the total uncertainty. Effectively, the proposed method provides a framework to obtain near-optimal solutions for the orbital circumnavigation problem by jointly considering control, planning and estimation. To evaluate the approach, we consider scenarios of a satellite circumnavigating another satellite in orbit while observing a set of feature points (landmarks) on the target satellite. The objective of the observing satellite is to accurately map the landmarks. Simulation results indicate that the proposed method results in trajectories which compare favorably to a heuristic "greedy" planning algorithm in terms of landmark uncertainty reduction.

It should be noted that Cross Entropy (CE) optimization has been used in the past for obstacle avoidance of Unmanned Aerial Vehicles (UAV), see Ref. [16]. Herein the CE method is applied for active localization in space with the intention of accurately localizing a set of landmarks whose location is known a-priori with high uncertainty. It is worth-mentioning that despite previous similar work of CE optimization for mobile robotic agents, the current paper is the first one to apply this approach to the relative pose localization problem in orbit, at least as far as the authors know. This problem also differs from previous work because in our particular case, most of the state variables of the observer satellite are constrained to follow a specific circular orbit. As a final remark, it should be mentioned that although information-theoretic methods (including CE optimization) are computationally expensive for most terrestrial and aerial applications (which require short time reaction scales), the long duration times encountered in orbit applications make the CE technique a feasible alternative for space applications.

The paper is organized as follows: in Section II we briefly introduce the problem to be solved along with the basic assumptions, definitions and kinematic equations. In Section III we provide a brief description the Cross Entropy method as it is applied to solve stochastic optimal control problems. In Section IV we present the state and observation models for the relative navigation scenario of two satellites along with the EKF equations used to estimate the state. Section V presents the results of our numerical simulation experiments, and we conclude this paper in Section VI with a summary of our work and some suggestions for future extensions.

II. Relative Navigation in Orbit

In this paper we consider the following problem: a main "target" satellite is orbiting the Earth in a circular or elliptical orbit. Using a (perhaps) smaller satellite (the "observer"), we wish to monitor/observe the target satellite closely using a single camera. The camera is assumed to have a restricted field-of-view and limited resolution range. In addition, the camera is allowed to gimbal and track "informative" features on the target satellite. The generation and determination of such features are typically the output of well-known image processing algorithms[17,18] and are not the subject in this study. For simplicity, it will be assumed that the camera is tracking the "most informative" feature at each instant of time, and that the total number N of available features is given a priori, although these assumptions can be easily removed. Finally, we assume that the vision processing pipeline deals with the (in general, non-trivial) data association and feature correspondence problems.[19–21]

As the observer/chaser satellite circumnavigates around the target satellite, it uses the tracked feature points on the target satellite to localize itself with respect to the latter. Localization here means determining the relative orientation of the observer satellite with respect to the target satellite. Since the observer satellite is in a predetermined orbit, it is assumed that the relative distance between the two is known to the observer satellite. Hence, only the relative orientation between the two satellites needs to be determined.

To show the feasibility of the proposed approach, we consider only a simplified/planar version of the previous problem. Specifically, we assume that the observer/chaser satellite circumnavigates the target satellite following a circular trajectory of radius R_{orb} having linear velocity V and relative orbital velocity $\omega_\phi = V/R_{\text{orb}}$. Note that the choice of a circular orbit is only made for the sake of simplicity. Typical relative orbits of two satellites flying in formation would result in a 2×1 elliptical orbit.[22] The incorporation of an elliptical reference orbit for the chaser satellite is straightforward.

In order to simplify the analysis, we will assume that the target satellite is stationary, and that the observer satellite (or sensor) is free to rotate around the axis that is normal to the xy plane going through its center of mass (the $z_{R;S}$ axis of the $\{R\}$ and $\{S\}$ frames in Figure 1) at an angle θ, which is independent of the angle ϕ.

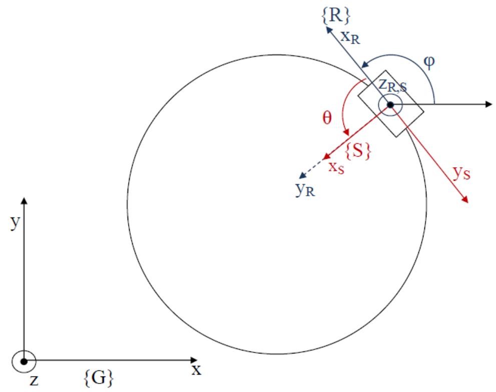


Figure 1. The various coordinate systems for an observer in circular orbit carrying a sensor: Global non Rotating Coordinate System {G}, Local non Rotating Coordinate System centered on the observer {R}, and Local Rotating Coordinate System on the sensor {S} .

The tracked features on the target satellite serve as "landmarks," whose location is to be determined and may denote feature points provided by a range and bearing sensor. Detection of the landmarks occurs only if they are within the field of view and range of the sensor, as shown in Figure 2.

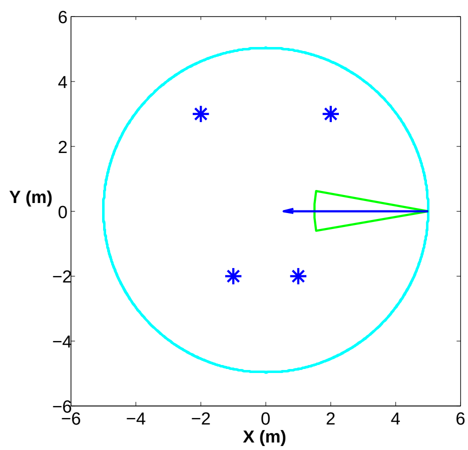


Figure 2. A sample simulation setup showing the landmarks (blue stars), the orbit of the observer (cyan), and the sensing area of the sensor (green).

Augmenting the positions of the N landmarks, $\mathbf{p}_i \in \mathbb{R}^2$, to the orbiting satellite state vector yields the vector $\mathbf{x} = (x, y, \phi, \theta, \mathbf{p}_1, ..., \mathbf{p}_N)^\mathsf{T} \in \mathbb{R}^{2N+4}$. The problem kinematics are then expressed as follows:

$$\begin{pmatrix} \mathrm{d}x(t) \\ \mathrm{d}y(t) \\ \mathrm{d}\phi(t) \\ \mathrm{d}\theta(t) \\ \mathrm{d}\mathbf{p}_1(t) \\ \vdots \\ \mathrm{d}\mathbf{p}_N(t) \end{pmatrix} = \begin{pmatrix} V\cos\phi(t)\mathrm{d}t \\ V\sin\phi(t)\mathrm{d}t \\ \omega_\phi(t)\mathrm{d}t \\ \omega_\theta(t)\mathrm{d}t \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \mathbf{I}_4 \\ \mathbf{0}_{2N\times 4} \end{pmatrix} \mathrm{d}\mathbf{w}(t), \tag{1}$$

where x, y correspond to the position, ϕ to the orientation of the observer satellite with respect to the global reference frame {G} and θ is the orientation of the sensor with respect to the local frame {R}. The control input is the angular velocity ω_θ. The process noise, $\mathrm{d}\mathbf{w}(t) \in \mathbb{R}^4$, is assumed to be white Gaussian with covariance matrix $\mathbf{\Sigma_w} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2)$.

The observation model can be written as follows:

$$\mathrm{d}\mathbf{y}(t) = {}^{\mathsf{S}}_{\mathsf{R}}\mathbf{R}(\theta(t)){}^{\mathsf{R}}_{\mathsf{G}}\mathbf{R}(\phi(t))\left(\mathbf{p}_i(t) - \mathbf{p}_R(t)\right)\mathrm{d}t + \mathrm{d}\mathbf{v}(t), \tag{2}$$

where $\mathbf{p}_i = (p_{x_i}, p_{y_i})^\mathsf{T}$ and $\mathbf{p}_R = (x, y)$ are the position of the landmarks and the location of the observer satellite, respectively. The term $\mathrm{d}\mathbf{v}$ corresponds to the observation noise of the sensor, which is considered to be zero-mean Gaussian with covariance matrix $\mathbf{\Sigma_v} = \mathrm{diag}(\sigma_5^2, \sigma_6^2)$. The matrices ${}^{\mathsf{R}}_{\mathsf{G}}\mathbf{R}(\phi(t_k))$ and ${}^{\mathsf{S}}_{\mathsf{R}}\mathbf{R}(\theta(t_k))$ express rotational transformations from the global $\{G\}$ to the observer frame of reference $\{R\}$ and from $\{R\}$ to the sensor reference frame $\{S\}$, respectively. In compact form, the observation model is written as

$$\mathrm{d}\mathbf{y}(t) = \mathbf{H}\left(\mathbf{x}(t)\right)\mathrm{d}t + \mathrm{d}\mathbf{v}(t). \tag{3}$$

As it was shown in Ref. [23], using a cost function that includes a measure of estimation uncertainty can lead to trajectories that reduce the overall uncertainty of a map. A good candidate for such a measure of uncertainty is the trace of the covariance matrix for the state estimate provided by an Extended Kalman Filter (EKF), using Eq. (1) and Eq. (2) for the propagation and update phase of the EKF, respectively. Alternatively, an Unscented Kalman Filter (UKF) can be used as described in Ref. [24]. The cost function is therefore of the form:

$$\boldsymbol{\mathcal{L}}(\mathbf{x}, \mathbf{u}) = \psi(\mathbf{x}_{t_N}) + \int_0^{t_N}\left(Q(\mathbf{x}) + \frac{1}{2}\mathbf{u}^\mathsf{T} R\mathbf{u}\right)\mathrm{d}t, \tag{4}$$

where the terminal cost is

$$\psi(\mathbf{x}_{t_N}) = \|\mathbf{p}^* - \mathbf{p}_R\|^2 + w_{\mathbf{\Sigma}}\,\mathrm{trace}\left(\mathbf{\Sigma}(t_N)\right), \tag{5}$$

with $\mathbf{p}^*$ being the target position, $\mathbf{p}_R$ the final position of the observer in the global reference frame $\{G\}$ and $w_{\mathbf{\Sigma}}$ the weight for the trace of the covariance given by the EKF at terminal time t_N.

III. Stochastic Optimal Control and Relative Entropy Optimization

In order to solve the stochastic minimization problem (1)-(4) we use ideas from Cross-Entropy optimization. The main ideas behind CE optimization are given in the next section. For more details, the interested reader is referred to Refs. [25] and [16].

III.A. Cross Entropy Formulation

In this section we summarize the main ideas of Cross Entropy minimization and how it can be used to solve a certain class of stochastic optimal control problems. To this end, assume that we are given a system with stochastic dynamics

$$\mathbf{dx} = \mathbf{f}(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{g}(\mathbf{x})\mathrm{d}\mathbf{w}, \tag{6}$$

and we wish to minimize the cost

$$\min_{\mathbf{u}} \mathbb{E}_p[\boldsymbol{\mathcal{L}}(\mathbf{x}, \mathbf{u})], \tag{7}$$

where the expectation in Eq. (7) is with respect to the trajectories of Eq. (6). In Equation (6) $\mathbf{x} \in \mathbb{R}^n$ is the state of the system, $\mathbf{u} \in \mathbb{R}^p$ is the control vector, and $\mathbf{w} \in \mathbb{R}^\ell$ is a zero-mean Gaussian process with covariance $\mathbf{\Sigma_w}$.

We parameterize the control input $\mathbf{u}(t)$ in terms of a parameter vector $\boldsymbol{\lambda} \in \mathbf{\Lambda} \subset \mathbb{R}^m$ to obtain $\mathbf{u}(t) = \mathbf{u}(t; \boldsymbol{\lambda})$. Therefore, instead of minimizing the cost function in Eq. (7) with respect to $\mathbf{u}(t)$ we will minimize it with respect to the finite dimensional vector $\boldsymbol{\lambda}$.

According to the Cross Entropy minimization method,[25] we rewrite the cost function as follows:

$$\mathcal{J}(\boldsymbol{\lambda}) = \mathbb{E}_p[\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})] = \int_{\mathbf{\Lambda}} p(\boldsymbol{\lambda})\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})\mathrm{d}\boldsymbol{\lambda}, \tag{8}$$

where $p(\boldsymbol{\lambda})$ is the probability density function corresponding to sampling trajectories based on Eq. (6). Under a parameterization of the baseline probability density, we have $p(\boldsymbol{\lambda}) = p(\boldsymbol{\lambda}; \boldsymbol{\mu})$. We perform importance sampling from a proposal probability density $q(\boldsymbol{\lambda})$ and evaluate the expectation in Eq. (8) as follows:

$$\mathcal{J}(\boldsymbol{\lambda}) = \int \frac{p(\boldsymbol{\lambda};\boldsymbol{\mu})}{q(\boldsymbol{\lambda})}\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})q(\boldsymbol{\lambda})\mathrm{d}\boldsymbol{\lambda} = \mathbb{E}_q\left[\frac{p(\boldsymbol{\lambda};\boldsymbol{\mu})}{q(\boldsymbol{\lambda})}\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})\right]. \tag{9}$$

The expression above can be approximated numerically by

$$\hat{\mathcal{J}}(\boldsymbol{\lambda}) \approx \frac{1}{N_s}\sum_{i=1}^{N_s}\left[\frac{p(\boldsymbol{\lambda}_i;\boldsymbol{\mu})}{q(\boldsymbol{\lambda}_i)}\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda}_i)\right] \tag{10}$$

with $\hat{\mathcal{J}}(\boldsymbol{\lambda})$ being an unbiased estimator and N_s the number of samples drawn. The probability density that minimizes the variance of the estimator $\hat{\mathcal{J}}(\boldsymbol{\lambda})$ is

$$q^*(\boldsymbol{\lambda}) = \underset{q}{\operatorname{argmin}}\operatorname{Var}\left[\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})\frac{p(\boldsymbol{\lambda};\boldsymbol{\mu})}{q(\boldsymbol{\lambda})}\right] = \frac{p(\boldsymbol{\lambda};\boldsymbol{\mu})\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})}{\mathcal{J}(\boldsymbol{\lambda})} \tag{11}$$

and it is the optimal (with respect to variance) importance sampling density.

In the cross entropy optimization framework[25] the main idea is to find the parameters $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ within the parametric class of pdfs $p(\boldsymbol{\lambda};\boldsymbol{\Xi})$, such that the probability distribution $p(\boldsymbol{\lambda};\boldsymbol{\xi})$ is close to the optimal distribution $q^*(\boldsymbol{\lambda})$ given in Eq. (11). Using the Kullback-Leibler divergence as a distance metric between $q^*(\boldsymbol{\lambda})$ and $p(\boldsymbol{\lambda};\boldsymbol{\xi})$ yields

$$\mathsf{D}\left(q^*(\boldsymbol{\lambda})\|p(\boldsymbol{\lambda};\boldsymbol{\xi})\right) = \int q^*(\boldsymbol{\lambda})\ln q^*(\boldsymbol{\lambda})\mathrm{d}\boldsymbol{\lambda} - \int q^*(\boldsymbol{\lambda})\ln p(\boldsymbol{\lambda};\boldsymbol{\xi})\mathrm{d}\boldsymbol{\lambda}.$$

The minimization problem can now be formulated as follows:

$$\begin{aligned}
\boldsymbol{\xi}^* &= \underset{\boldsymbol{\xi}}{\operatorname{argmin}}\,\mathsf{D}\left(q^*(\boldsymbol{\lambda})\|p(\boldsymbol{\lambda};\boldsymbol{\xi})\right)\\
&= \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\int q^*(\boldsymbol{\lambda})\ln p(\boldsymbol{\lambda};\boldsymbol{\xi})\mathrm{d}\boldsymbol{\lambda}\\
&= \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\int \frac{p(\boldsymbol{\lambda};\boldsymbol{\mu})\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})}{\mathcal{J}(\boldsymbol{\lambda})}\ln p(\boldsymbol{\lambda};\boldsymbol{\xi})\mathrm{d}\boldsymbol{\lambda}\\
&= \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\int p(\boldsymbol{\lambda};\boldsymbol{\mu})\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})\ln p(\boldsymbol{\lambda};\boldsymbol{\xi})\mathrm{d}\boldsymbol{\lambda}\\
&= \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\,\mathbb{E}_{p(\boldsymbol{\lambda};\boldsymbol{\mu})}\left[\boldsymbol{\mathcal{L}}(\mathbf{x})\ln p(\boldsymbol{\lambda};\boldsymbol{\xi})\right].
\end{aligned}$$

Based on the previous equation, the optimal parameters can be approximated numerically by

$$\boldsymbol{\xi}^* \approx \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\frac{1}{N_s}\sum_{i=1}^{N_s}\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda}_i)\ln p(\boldsymbol{\lambda}_i;\boldsymbol{\xi}) \tag{12}$$

with N_s sample paths evaluated under the density $p(\boldsymbol{\lambda};\boldsymbol{\mu})$.

III.B. Optimization as Estimation of Rare-Event Probabilities

This subsection demonstrates how the CE optimization recasts the minimization of the cost shown in Eq. (8) as an *estimation problem of rare event probability.* To this end, we consider the problem of estimating the probability that a trajectory $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{t_N}]$, which arose by sampling control vector parameters from the distribution $p(\boldsymbol{\lambda};\boldsymbol{\mu})$, has a cost that is smaller than a constant γ. Specifically, we have

$$\mathbb{P}\left(\boldsymbol{\mathcal{L}} \le \gamma\right) = \mathbb{E}_{p(\boldsymbol{\lambda};\boldsymbol{\mu})}[\mathrm{I}_{\{\boldsymbol{\mathcal{L}}\le\gamma\}}].$$

Using (10), this probability can be numerically approximated by

$$\hat{\mathbb{P}}\left(\boldsymbol{\mathcal{L}} \le \gamma\right) \approx \frac{1}{N_s}\sum_{i=1}^{N_s}\left[\frac{p(\boldsymbol{\lambda}_i;\boldsymbol{\mu})}{p(\boldsymbol{\lambda}_i;\boldsymbol{\xi})}\mathrm{I}_{\{\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda}_i)\le\gamma\}}\right],$$

where $\boldsymbol{\lambda}_i$ are i.i.d samples drawn from $p(\boldsymbol{\lambda};\boldsymbol{\xi})$. Based on Eq. (12), the goal becomes to find the optimal $\boldsymbol{\xi}^*$ which is defined as

$$\boldsymbol{\xi}^* = \underset{\boldsymbol{\xi}}{\operatorname{argmax}} \frac{1}{N_s} \sum_{i=1}^{N_s} \mathrm{I}_{\{\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda}_i) \leq \gamma\}} \ln p(\boldsymbol{\lambda}_i;\boldsymbol{\xi}), \tag{13}$$

where now the samples $\boldsymbol{\lambda}_i$ are generated according to probability density $p(\boldsymbol{\lambda};\boldsymbol{\mu})$. Since the event $\{\boldsymbol{\mathcal{L}} \leq \gamma\}$ is rare, its probability is difficult to be estimated directly using, for example, brute force Monte-Carlo. Instead, one may start with a $\gamma_1 > \gamma$ for which the probability of the event $\{\boldsymbol{\mathcal{L}} \leq \gamma_1\}$ is equal to some $\rho > 0$. Thus, the value γ_1 is set to the ρ-th quantile of $\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})$ which means that γ_1 is the largest number for which

$$\mathbb{P}\left(\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda}) \leq \gamma_1\right) = \rho. \tag{14}$$

The parameter γ_1 can be found by sorting the samples according to their cost in increasing order and setting $\gamma_1 = \boldsymbol{\mathcal{L}}_{\lceil \rho N \rceil}$. The optimal parameter $\boldsymbol{\xi}_1$ for the γ_1 level is calculated according to Eq. (13) using the value γ_1. This iterative procedure terminates when $\gamma_k \leq \gamma$, in which case the corresponding parameter $\boldsymbol{\xi}_k$ is the optimal one and thus $\boldsymbol{\xi}^* = \boldsymbol{\xi}_k$.

To summarize, for the case of optimal control in continuous action spaces, optimization is treated as an *estimation problem of rare event probability.* To find the optimal trajectory $\boldsymbol{\lambda}^*$ and optimal parameters $\boldsymbol{\xi}^*$, the process of estimating rare event probabilities is iterated until $\gamma \to \gamma^*$, where $\gamma^* = \min \boldsymbol{\mathcal{L}}(\boldsymbol{\lambda})$. Since γ^* is not known *a-priori*, we choose as γ^* the value of γ for which no further improvement in the iterative process is observed.

Iterative Algorithm. Choosing $p(\boldsymbol{\lambda};\boldsymbol{\xi})$ to be a Gaussian, the algorithm can be summarized as follows:

1. Initialize an iteration counter $j \leftarrow 1$ and select some initial parameters $\boldsymbol{\mu} = \boldsymbol{\xi}_1 = \{\mu_{\boldsymbol{\xi}_1}, \Sigma_{\boldsymbol{\xi}_1}\}$.
2. From a pdf $p(\boldsymbol{\lambda};\boldsymbol{\xi}_j) = \boldsymbol{\mathcal{N}}(\mu_{\boldsymbol{\xi}_j}, \Sigma_{\boldsymbol{\xi}_j})$ draw a set of N_s parameter vectors $\boldsymbol{\lambda}_i$.
3. Evaluate the cost of each trajectory $J(\boldsymbol{\lambda}_i)$ and sort them in ascending order.
4. Select the best performing ρ-th percentile to form a set $\boldsymbol{\mathcal{S}}_j$ and use it to estimate the parameters $\boldsymbol{\xi}_j$:

$$\mu_{\boldsymbol{\xi}_j} = \frac{1}{|\boldsymbol{\mathcal{S}}_j|} \sum_{\vec{\mathbf{x}}_i \in \boldsymbol{\mathcal{S}}_j} \boldsymbol{\lambda}_i, \tag{15}$$

$$\Sigma_{\boldsymbol{\xi}_j} = \frac{1}{|\boldsymbol{\mathcal{S}}_j|} \sum_{\vec{\mathbf{x}}_i \in \boldsymbol{\mathcal{S}}_j} (\boldsymbol{\lambda}_i - \mu_{\boldsymbol{\xi}_j})(\boldsymbol{\lambda}_i - \mu_{\boldsymbol{\xi}_j})^{\mathsf{T}} \tag{16}$$

5. If the stopping criterion is met set $\boldsymbol{\xi}^* \leftarrow \boldsymbol{\xi}_j$ and halt. Otherwise, set $j \leftarrow j+1$ and iterate from step 2 using the newly adjusted parameters $\boldsymbol{\xi}_j$.

IV. Application of CE Optimization to Satellite Self-Localization in Orbit

In order to apply the results of the previous section to our problem, we first consider a discretized version of Eq. (1) as follows:

$$\begin{pmatrix} x(t_{k+1}) \\ y(t_{k+1}) \\ \phi(t_{k+1}) \\ \theta(t_{k+1}) \\ \mathbf{p}_1(t_{k+1}) \\ \dots \\ \mathbf{p}_N(t_{k+1}) \end{pmatrix} = \begin{pmatrix} x(t_k) + V\cos(\phi(t_k))\delta t \\ y(t_k) + V\sin(\phi(t_k))\delta t \\ \phi(t_k) + \omega_\phi(t_k)\delta t \\ \theta(t_k) + \omega_\theta(t_k)\delta t \\ \mathbf{p}_1(t_k) \\ \dots \\ \mathbf{p}_N(t_k) \end{pmatrix} + \begin{pmatrix} \mathbf{I_4} \\ \mathbf{0_{2N\times 4}} \end{pmatrix} \mathbf{w}(t_k), \tag{17}$$

where δt is the discretization step. We use an EKF to estimate the state of the system, given the measurements. The propagation phase of an EKF for the system described in Eq. (17) is then given by[7]

$$\begin{aligned}\hat{\mathbf{x}}(t_{k+1|k}) &= \mathbf{f}\Big(\hat{\mathbf{x}}(t_{k|k}), \mathbf{u}(\hat{\mathbf{x}}(t_{k|k}), t_k; \boldsymbol{\lambda})\Big)\\ \boldsymbol{\Phi} &= \nabla_{\hat{\mathbf{x}}}\mathbf{f}\Big(\hat{\mathbf{x}}(t_{k|k}), \mathbf{u}(\hat{\mathbf{x}}(t_{k|k}), t_k; \boldsymbol{\lambda})\Big)\\ \boldsymbol{\Sigma}(t_{k+1|k}) &= \boldsymbol{\Phi}\boldsymbol{\Sigma}(t_{k|k})\boldsymbol{\Phi}^{\mathsf{T}} + \boldsymbol{\Sigma}_{\mathbf{w}},\end{aligned} \tag{18}$$

where $\boldsymbol{\lambda}$ is a parameter vector indicating that the control is parameterized. Similarly, the observation model in Eq. (2) can be written in a discrete form as

$$^{\mathsf{S}}\mathbf{y}(t_k) = {}^{\mathsf{S}}_{\mathsf{R}}\mathbf{R}(\theta(t_k)){}^{\mathsf{R}}_{\mathsf{G}}\mathbf{R}(\phi(t_k))\Big({}^{\mathsf{G}}\mathbf{p}_i(t_k) - {}^{\mathsf{G}}\mathbf{p}_{\mathsf{R}}(t_k)\Big) + \mathbf{v}(t_k). \tag{19}$$

In a compact form the discrete observation model is given by

$$\mathbf{y}(t_k) = \mathbf{H}\left(\mathbf{x}(t_k)\right) + \mathbf{v}(t_k). \tag{20}$$

The function $\mathbf{H}\left(\mathbf{x}(t_k)\right)$ is defined such that it matches the observation scenario as expressed in Eq. (19). Given the observation model, the Extended Kalman Filter update equations are:

$$\begin{aligned}\boldsymbol{\mathcal{H}}(\hat{\mathbf{x}}(t_{k+1|k})) &= \nabla_{\hat{\mathbf{x}}(t_{k+1|k})}\mathbf{H}(\mathbf{x}(t))\\ \hat{\mathbf{y}}(t_{k+1}) &= \mathbf{H}(\hat{\mathbf{x}}(t_{k+1|k}))\\ \mathbf{y}(t_{k+1}) &= \mathbf{H}(\mathbf{x}(t_{k+1})) + \mathbf{v}(t_{k+1})\\ \mathbf{r}(t_{k+1}) &= \mathbf{y}(t_{k+1}) - \hat{\mathbf{y}}(t_{k+1})\\ \boldsymbol{\mathcal{S}}(t_{k+1}) &= \boldsymbol{\mathcal{H}}\boldsymbol{\Sigma}(t_{k+1|k})\boldsymbol{\mathcal{H}}^{\mathsf{T}} + \boldsymbol{\Sigma}_{\mathbf{v}}\\ \boldsymbol{\mathcal{K}}(t_{k+1}) &= \boldsymbol{\Sigma}(t_{k+1|k})\boldsymbol{\mathcal{H}}^{\mathsf{T}}\boldsymbol{\mathcal{S}}(t_{k+1})^{-1}\\ \hat{\mathbf{x}}(t_{k+1|k+1}) &= \hat{\mathbf{x}}(t_{k+1|k}) + \boldsymbol{\mathcal{K}}(t_{k+1})\mathbf{r}(t_{k+1})\\ \boldsymbol{\Sigma}(t_{k+1|k+1}) &= \left(I - \boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{H}}\right)\boldsymbol{\Sigma}(t_{k+1|k})\left(I - \boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{H}}\right)^{\mathsf{T}} + \boldsymbol{\mathcal{K}}\boldsymbol{\Sigma}_{\mathbf{v}}\boldsymbol{\mathcal{K}}^{\mathsf{T}}.\end{aligned} \tag{21}$$

Since the sensor has a limited field of view, the choice of ω_θ may have a significant influence on the uncertainty of the estimate of the state $\mathbf{x}(t_N)$ as it is captured by the trace($\boldsymbol{\Sigma}(t_N)$), where t_N is the time horizon of a trajectory. We therefore choose ω_θ such that it minimizes the descrete cost:

$$\hat{\boldsymbol{\mathcal{L}}}(\mathbf{x}, \mathbf{u}) \approx \psi(\mathbf{x}_{t_N}) + \sum_{k=0}^{N}\left(\frac{1}{2}\mathbf{u}(t_k)^{\mathsf{T}}R\mathbf{u}(t_k)\right)\delta t. \tag{22}$$

For simplicity, we let $Q(\mathbf{x}) = 0$ and $\|\mathbf{p}^* - \mathbf{p}_R\| = 0$. We use the following parameterization of ω_θ,

$$\begin{aligned}\omega_\theta(t_k) &= \mathbf{u}\Big(\omega_\theta(t_{k-1}), \alpha_{\omega_\theta}(t_{k-1}; \boldsymbol{\lambda})\Big)\\ &= \omega_\theta(t_{k-1}) + \alpha_{\omega_\theta}(t_{k-1}; \boldsymbol{\lambda})\delta t,\end{aligned} \tag{23}$$

where $\alpha_{\omega_\theta}(t_{k-1}; \boldsymbol{\lambda})$ is the rotational acceleration, which is parameterized as a piecewise constant trajectory split in M parts. For each part i there are two parameters, the duration ΔT_i and the angular acceleration $\alpha_i \in [\alpha_{\min},\ \alpha_{\max}] \subset \mathbb{R}$ which remains constant in time interval ΔT_i. The interval $[\alpha_{\min}, \alpha_{\max}]$ defines the range of acceptable angular acceleration values. The sum of all time intervals is fixed and equal to the pre-specified time horizon, that is, $\sum_{i=1}^{M} \Delta T_i = t_N$. The parameter vector $\boldsymbol{\lambda}$ is, therefore, defined as $\boldsymbol{\lambda}^{\mathsf{T}} = (\Delta T_1, \alpha_1, ..., \Delta T_M, \alpha_M)$. Each parameter vector $\boldsymbol{\lambda}$ corresponds to a unique control vector $\mathbf{u}$, which, given an initial state value $\mathbf{x}_0$, gives rise to a unique trajectory $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{t_N}]$. Formally, one can say that there is a function $f_{\boldsymbol{\lambda}} : \boldsymbol{\Lambda} \rightarrow \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{U}}$ where $\boldsymbol{\Lambda}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{U}}$ denote the sets of the parameter, state and control trajectories respectively. Given this parameterization, the cost function in (4) can be restated as

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\lambda}) = \boldsymbol{\mathcal{C}}(f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda})). \tag{24}$$

V. Numerical Results

To evaluate the proposed method, we perform a series of numerical simulations. We assume an observer satellite at a distance of approximately 6 meters from the target satellite. Given the close distance and the fact that the camera has a limited field of view, the satellite needs to actively move the camera to be able to detect all available landmarks on the target satellite.

In the first simulation scenario, we compare the uncertainty measure for the state of the observer and the position of the landmarks between the cases when $\omega_\theta = 0$ and when it is provided by the proposed CE optimization method. As can be seen in Figure 3, the estimator remains consistent throughout the duration of the trajectory. For all cases, the optimization horizon t_N is taken to be equal to one complete orbit, assumed for our scenario to to take about 30 minutes.

Figure 4 shows the 3σ uncertainty ellipses around the landmarks at the time horizon from which it becomes evident that the bounds are tighter for the optimized case. As an additional indication of this, at the end of the time horizon, the value of trace($\mathbf{\Sigma}_{t_N}$) was 12.15 for the case where optimization was used and 14.56 for the one without implying lower uncertainty for the optimized case.

The main observation from these results is that the optimized ω_θ produces trajectories which place the landmarks within the sensing area for longer periods of time, so that better localization is achieved. An example of such behavior is shown in Figure 5.

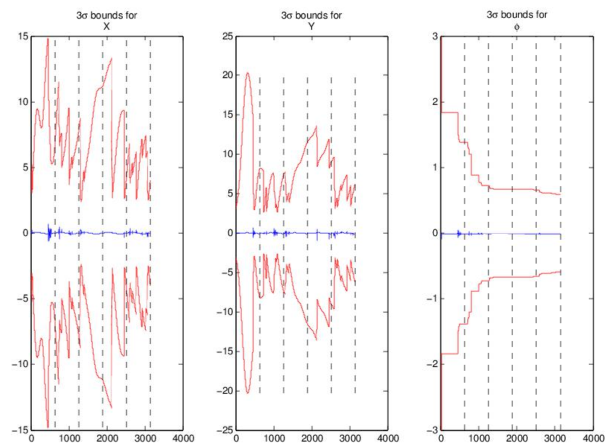


Figure 3. Estimation error (in blue) and 3σ bounds (in red) for the state variables of the observer over time when ω_θ is given through the Relative Entropy optimization. Notice that the estimator remains consistent.

Next, we compare the proposed method against a heuristic "greedy" algorithm that uses a simple linear controller, designed to align the observer with the nearest landmark. Specifically, the rotational velocity of the observer, $\omega_\theta^{\text{greedy}}$, is expressed as

$$\begin{aligned}\omega_\theta^{\text{greedy}}(t_k) &= \mathbf{u}\Big(\omega_\theta^{\text{greedy}}(t_{k-1}), \alpha_{\omega_\theta}^{\text{greedy}}(t_{k-1})\Big)\\ &= \omega_\theta^{\text{greedy}}(t_{k-1}) + K_p \min_i \|\mathbf{p}_i - \mathbf{p}_R(t_{k-1})\|^2 \delta t,\end{aligned} \tag{25}$$

where K_p is the proportional gain, $\mathbf{p}_R(t_{k-1})$ the position of the observer at time instance t_{k-1} and $\mathbf{p}_i$ the position of landmark i. To make the comparison as fair as possible, the allowable angular acceleration interval $[\alpha_{\min}, \alpha_{\max}]$ was kept the same as the one in the parameterization of the control policy used in the CE algorithm. As can be seen in Figure 6, the 3σ bounds for the location of the landmarks produced by the CE algorithm are similar to those of the "greedy" for landmarks 3 and 4. The greedy algorithm results in slightly tighter bounds for landmarks 1 and 2, at the expense of the bounds for landmark 5, where it is significantly outperformed by the proposed method. Overall, the measure of uncertainty for the trajectory selected by the CE method is smaller than the one resulting from the "greedy", trace $(\mathbf{\Sigma}_{\text{CE}}(t_N)) = 14.70 < 15.26 = $ trace $(\mathbf{\Sigma}_{\text{greedy}}(t_N))$, indicating a higher uncertainty reduction for the entire landmark map. Taking into account the control cost incurred for each method ($\sum_{k=0}^{t_N} u_{\text{CE}}(t_k) R u_{\text{CE}}^\mathsf{T}(t_k) = 34.52 < 73.11 = \sum_{k=0}^{t_N} u_{\text{greedy}}(t_k) R u_{\text{greedy}}^\mathsf{T}(t_k)$) it becomes clear that the proposed method has a definite advantage over the greedy heuristic. The latter, can be made to achieve similar uncertainty reduction to the CE optimization,

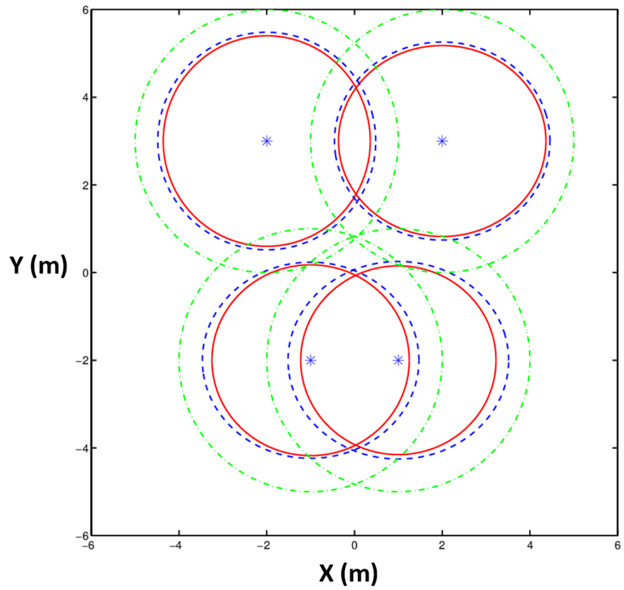


Figure 4. Uncertainty ellipses for the locations of the landmarks showing the 3σ bounds: (a) initially (green), (b) after two orbits where $\omega_\theta = 0$ (blue) and when ω_θ is given through the Relative Entropy optimization (red).

as shown in Figure 7, by increasing the control gain K_p and allowing for higher angular acceleration limits. However, the control cost in this case is at least three times higher that the one incurred using the proposed approach.

Finally, we briefly investigate the effect that the choice of the estimator (EKF vs UKF[24]) has on the estimated map uncertainty. The results, shown in Figure 8, indicate that the UKF yields tighter uncertainty bounds than the EKF. However, as expected, the UKF is computationally more expensive.

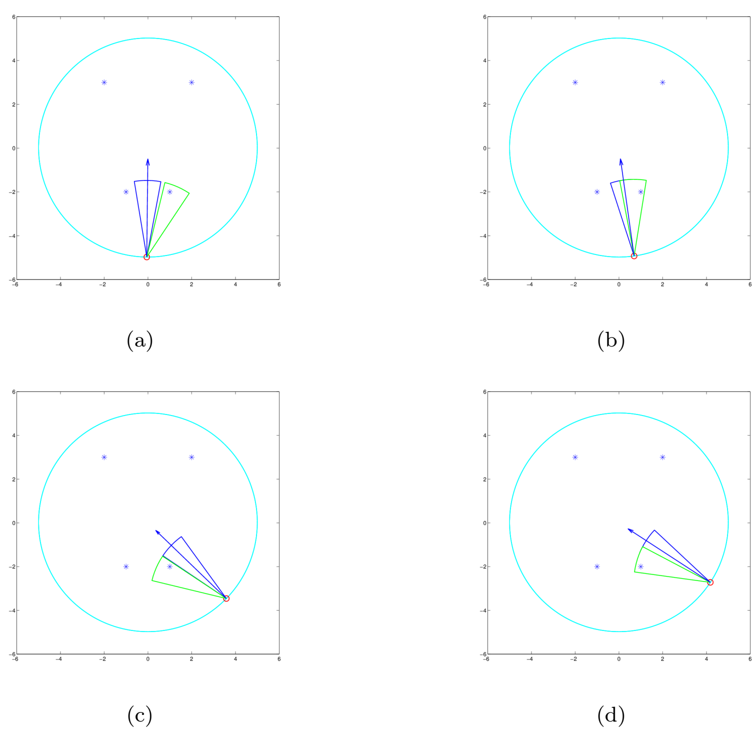


Figure 5. Examples of comparison of sensing instances between $\omega_\theta = 0$ (blue) and when ω_θ is given through the Relative Entropy optimization (green). In the latter case the landmark is kept within the sensing area for a longer period of time.

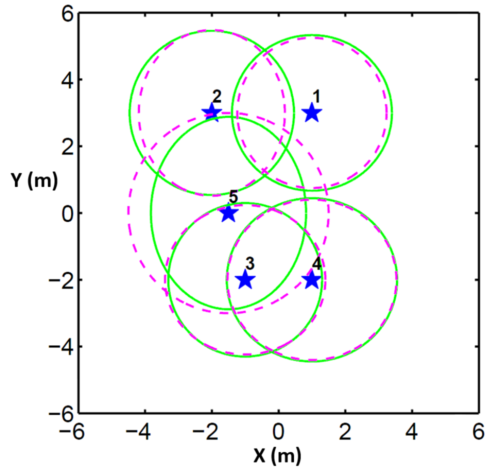


Figure 6. Comparison of 3σ uncertainty ellipses for the locations of the landmarks between a heuristic "greedy" algorithm (dashed magenta) and the Cross Entropy optimization (solid green).

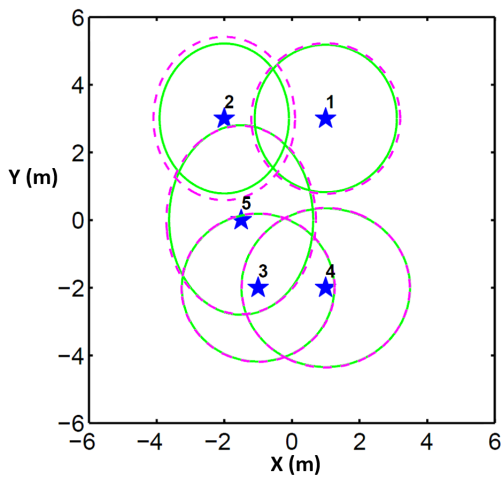
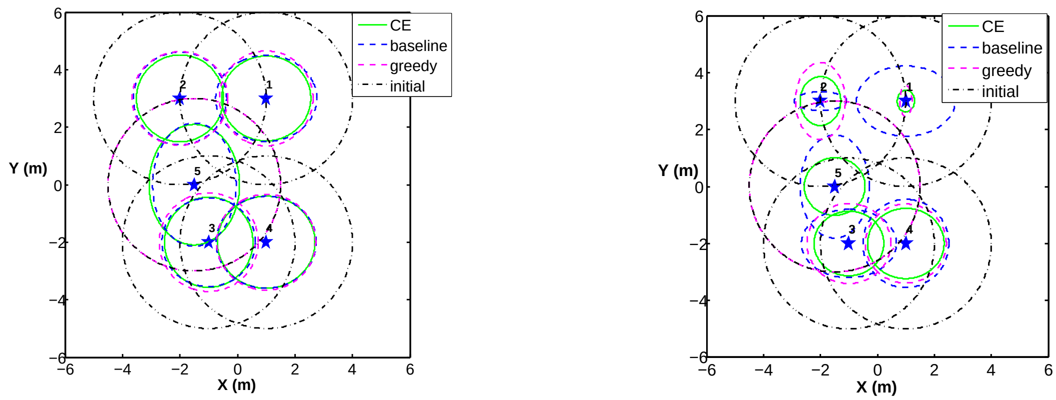


Figure 7. Comparison of 3σ uncertainty ellipses for the locations of the landmarks between a heuristic "greedy" algorithm (dashed magenta) and the Cross Entropy optimization (solid green). In this case the "greedy" essentially matches the uncertainty reduction of the proposed method at the expense of increased control cost.

VI. Conclusions

This paper presents a new approach to solving the active self-localization problem during relative navigation in orbit using Cross Entropy (CE) minimization. The method incorporates an uncertainty measure in the cost function, which can be utilized to select near-optimal trajectories in terms of estimation uncertainty. The proposed method, therefore, provides a principled way to balance control effort against localization uncertainty by jointly considering the planning, control and estimation problems. Numerical simulations comparing the results from the proposed approach against an open-loop strategy and a greedy strategy that simply tries to keep the landmarks visible throughout the orbit, show the benefits of the CE approach.

There are several potential improvements and extensions to the proposed approach. An immediate objective to consider, of course, is the use of more realistic orbital trajectories that account for the complete orbital dynamics of two satellites flying close to one another, such as the Cholessy-Wiltshire model.[22] This will allow the representation and analysis of more complex scenarios. This is part of on-going investigation. The major drawback of the CE approach is that is computationally expensive. Although the scale of the orbital dynamics is such that computing the optimal trajectories can be done in real-time, further improvements are possible. For example, the CE formulation is amenable to parallelization, which can be exploited to speed-up the solution.



(a) EKF (b) UKF

Figure 8. Comparison of 3σ uncertainty ellipses for the locations of the landmarks between a EKF (a) and an UKF (b) when used in conjunction with a heuristic "greedy" algorithm (dashed magenta) and the Cross Entropy optimization (solid green). The initial uncertainty bounds (in black) and the ones obtained by not actuating the sensor (blue baseline) are also shown.

Acknowledgments

This work is partially supported by AFRL grant FA9453-13-C-0201 and NSF award NRI-1426945.

References

[1]Ambrose, R., Wilcox, B., Reed, B., Matthies, L., Lavery, D., and Korsmeyer, D., "Robotics, tele-robotics, and autonomous systems roadmap," *NASAs Space Technology Roadmaps, National Aeronautics and Space Administration*, 2012.

[2]Augenstein, S., *Monocular Pose and Shape Estimation of Moving Targets, for Autonomous Rendezvous and Docking*, PhD thesis, Stanford University, 2011.

[3]Junkins, J. L., Hughes, D. C., Wazni, K. P., and Pariyapong, V., "Vision-Based Navigation For Rendezvous, Docking And Proximity Operations," *In AAS Guidance and Control Conference*, 1999, pp. 3–7.

[4]Kelsey, J. M., Byrne, J., Cosgrove, M., Seereeram, S., and Mehra, R. K., "Vision-based relative pose estimation for autonomous rendezvous and docking," *IEEE Aerospace Conference*, March 4–11, Big Sky, MT, 2006, pp. 20–pp.

[5]Singla, P., Subbarao, K., and Junkins, J. L., "Adaptive output feedback control for spacecraft rendezvous and docking under measurement uncertainty," *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 4, 2006, pp. 892–902.

[6]Tweddle, B. E., *Computer Vision Based Navigation for Spacecraft Proximity Operations*, Ph.D. thesis, Massachusetts Institute of Technology, 2010.

[7]Stengel, R. F., *Optimal Control and Estimation*, Dover books on Advanced Mathematics, Dover Publications, New York, 1994.

[8]Fleming, W. H. and Soner, H. M., *Controlled Markov Processes and Viscosity Solutions*, Applications of Mathematics, Springer, New York, 2nd ed., 2006.

[9]Dorato, P., Cerone, V., and Abdallah, C., *Linear Quadratic Control: An Introduction*, Krieger Publishing Co., Inc., Melbourne, FL, 2000.

[10]Sim, R. and Roy, N., "Active Exploration Planning for SLAM using Extended Information Filters," *In Proc. 20th Conf. Uncertainty in AI*, 2004.

[11]Leung, C., Huang, S., and Dissanayake, G., "Active SLAM in structured environments." *IEEE International Conference on Robotics and Automation* , May 19–23, Pasadena, CA, 2008, pp. 1898–1903.

[12]Leung, C., Huang, S., and Dissanayake, G., "Active SLAM using Model Predictive Control and Attractor based Exploration," *IEEE/RSJ International Conference on Intelligent Robots and Systems, October 9-15, Beijing, China*, 2006, pp. 5026–5031.

[13]Choudhary, S., Trevor, A. J., Christensen, H. I., and Dellaert, F., "SLAM with Object Discovery, Modeling and Mapping," *IEEE/RSJ International Conference on Intelligent Robots and Systems* , Sept. 14–18, Chicago, IL, 2014, pp. 1018–1025.

[14]Dong, J., Nelson, E., Indelman, V., Michael, N., and Dellaert, F., "Distributed Real-time Cooperative Localization and Mapping using an Uncertainty-Aware Expectation Maximization Approach," *IEEE International Conference on Robotics and Automation*, May 26–30, Seattle, WA, 2015.

[15]Choudhary, S., Indelman, V., Christensen, H. I., and Dellaert, F., "Information-based Reduced Landmark SLAM," *IEEE International Conference on Robotics and Automation*, May 26–30, Seattle, WA, 2015.

[16]Kobilarov, M., "Cross-Entropy Motion Planning," *The International Journal of Robotics Research*, Vol. 31, No. 37, pp. 855–871.

[17]Lowe, D. G., "Object Recognition from Local Scale-Invariant Features," *The Proceedings of the 7th IEEE International Conference on Computer Vision*, Vol. 2, Sept. 20–27, Kerkyra, Greece, 1999, pp. 1150–1157.

[18]Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L., "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, Vol. 110, No. 3, 2008, pp. 346–359.

[19]Neira, J. and Tardós, J. D., "Data association in stochastic mapping using the joint compatibility test," *IEEE Transactions on Robotics and Automation,*, Vol. 17, No. 6, 2001, pp. 890–897.

[20]Zhang, G., Kontitsis, M., Filipe, N., Tsiotras, P., and Vela, P. A., "Cooperative Relative Navigation for Space Rendezvous and Proximity Operations using Controlled Active Vision," *Journal of Field Robotics*, 2015, to appear.

[21]Collins, R. T., "Multitarget data association with higher-order motion models," *IEEE Conference on Computer Vision and Pattern Recognition*, June 16–21, Providence RI, 2012, pp. 1744–1751.

[22]Clohessy, H. and Wiltshire, R., "Terminal Guidance System for Satellite Rendezvous," *Journal of the Aerospace Sciences*, Vol. 27, No. 9, 1960, pp. 653–658.

[23]Kontitsis, M., Theodorou, E. A., and Todorov, E., "Multi-robot active slam with relative entropy optimization," *American Control Conference*, June 17–19, Washington, DC, 2013, pp. 2757–2764.

[24]Wan, E. A. and Van Der Merwe, R., "The Unscented Kalman Filter for Nonlinear Estimation," *The IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 2000, pp. 153–158.

[25]Botev, Z., Kroese, D., Rubinstein, R., and L'Ecuyer, P., "The Cross-Entropy Method for Optimization," Vol. 31 *Machine Learning*, In Handbook of Statistics, 2011.