

# Collaborative Pose Estimation of an Unknown Target Using Multiple Spacecraft

**Kai Matsuka**  
California Institute of Technology  
1200 E California Blvd.  
Pasadena, CA 91125, USA  
kmatsuka@caltech.edu

**Angel Santamaria-Navarro**  
Jet Propulsion Laboratory  
California Institute of Technology  
4800 Oak Grove Dr.  
Pasadena, CA 91109, USA  
angel.santamaria.navarro@jpl.nasa.gov

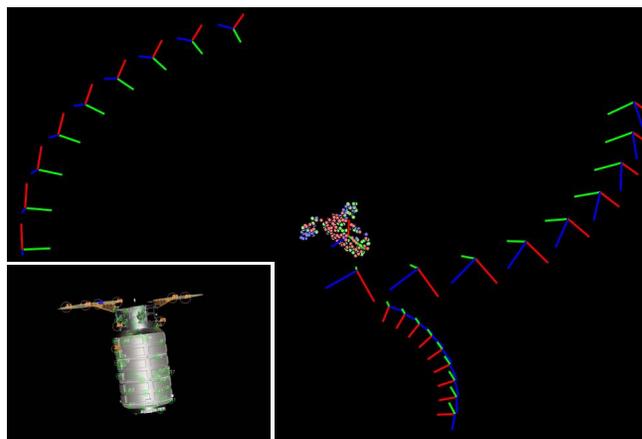
**Vincenzo Capuano**  
California Institute of Technology  
1200 E California Blvd.  
Pasadena, CA 91125, USA  
vcapuano@caltech.edu

**Alexei Harvard**  
California Institute of Technology  
1200 E California Blvd.  
Pasadena, CA 91125, USA  
alexh@caltech.edu

**Amir Rahmani**  
Jet Propulsion Laboratory  
California Institute of Technology  
4800 Oak Grove Dr.  
Pasadena, CA 91109, USA  
amir.rahmani@jpl.nasa.gov

**Soon-Jo Chung**  
California Institute of Technology  
1200 E California Blvd.  
Pasadena, CA 91125, USA  
sjchung@caltech.edu

**Abstract**—A reliable method for estimating the pose of an unknown and uncooperative space target using monocular vision remains an open problem. Vision-based pose determination is challenging due to factors such as harsh lighting conditions, rotational dynamics of the target, and scale ambiguity of the monocular camera. To address these challenges, we propose a novel collaborative pose determination algorithm called Multi-Spacecraft Simultaneous Estimation of Pose and Shape algorithm or M-SEPS. Within M-SEPS, a team of chaser spacecraft, each equipped with a monocular camera, exchange information over a local network to jointly estimate the relative kinematic state of the target and its sparse shape landmarks. In this approach, each spacecraft processes its images and extracts its own set of visual keypoints in parallel. Then, the team uses the local network to jointly estimate the target pose and shape in a distributed fashion by applying the consensus algorithm over the inter-spacecraft communication links. We validate our algorithm using simulations of relative orbits and observations captured by each chaser spacecraft. To the best of the authors’ knowledge, this is the first cooperative vision-based algorithm for estimating the pose and shape of a space object by means of an arbitrary number of spacecraft.



**Figure 1.** Target shape reconstructed from three cooperative spacecraft with monocular cameras. The bottom left image shows the keypoints seen by one of the chasers.

## TABLE OF CONTENTS

1. INTRODUCTION.....	1
2. PROBLEM STATEMENT .....	3
3. M-SEPS ARCHITECTURE OVERVIEW .....	4
4. EXTENDED DECENTRALIZED INFORMATION FILTER .....	5
5. SIMULATION.....	7
6. CONCLUSIONS.....	8
ACKNOWLEDGMENTS .....	9
REFERENCES .....	9
BIOGRAPHY .....	10

## 1. INTRODUCTION

Precise relative navigation solutions are essential to many space missions that involve formation flying and close-proximity operation. For example, the Space Shuttle rendezvous with the Mir Space Station, the Hubble Space Telescope, and the ISS [1] relied on the Trajectory Control Sensor and a scanning LIDAR; the Demonstration of Autonomous Rendezvous and Docking mission [2] utilized a state-of-the-art GPS relative navigation techniques and the Advanced Video Guidance Sensor [3]; the Orbital Express mission [4] adopted two cameras operating in the visible, one in the infrared spectrum, a laser range finder, as well as the Advanced Video Guidance Sensor already used for DART mission [5]; the European Space Agency’s Automated Transfer Vehicle missions employed the Videometer [6] and retro-reflectors to approach the ISS. More recently, RemoveDEBRIS [7], different navigation algorithms were validated in

978-1-7281-7436-5/21/\$31.00 ©2021 IEEE

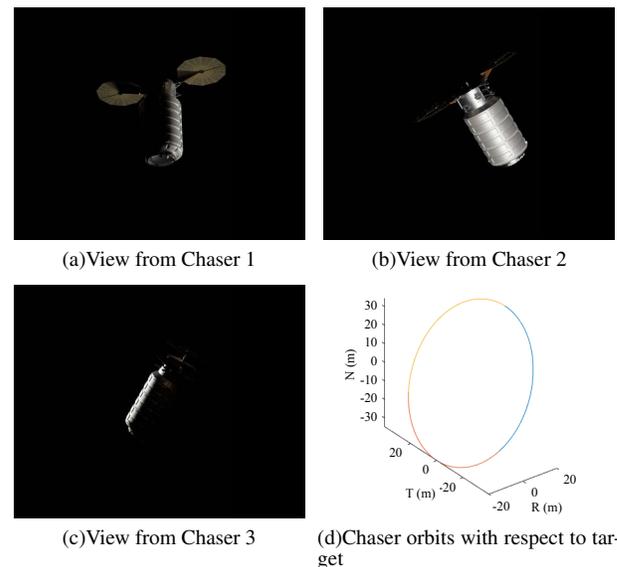
flight using an electro-optical-based navigation system based on both LIDAR and visual cameras together with differential GPS. The CubeSat Seeker [8], deployed from the spacecraft Cygnus, demonstrated the basic capabilities required for on-orbit inspection.

One of the relative navigation technologies that have gained attention recently in the literature is the vision-based relative pose tracking of artificial or natural objects in space. Vision-based pose estimation in close proximity has applications to on-orbit servicing, orbital debris removal, and exploration of small bodies [9]. Cameras are a popular relative sensor for small spacecraft as they have relatively low mass and power consumption, and a relatively high technology readiness level. However, vision-based algorithms have some challenges, such as highly variable and harsh lighting conditions in space. Moreover, the algorithms' approaches and difficulty are different depending on whether the target is *cooperative* or *uncooperative* and *known* or *unknown*. The estimation of *known* spacecraft leverages the known target model parameters such as geometry or inertia matrix, which simplifies the algorithm formulation. The less studied case is the problem of estimating relative motion for *uncooperative* and *unknown* orbiting objects, which is the focus of this paper. The development of a reliable vision-based relative pose tracking of an unknown, uncooperative target is still considered an open problem.

Due to the inaccessibility of space, along with the aforementioned challenges, only a limited number of flight experiments have been done related to vision-based pose estimation of another target. The SPHERES VERTIGO experiment [10] used robotic platforms inside the well-illuminated International Space Station (ISS) to demonstrate visual inspection and navigation, as well as 3D reconstruction of a free-flying target. PRISMA mission used images taken formation flying spacecraft in relative orbit, and the images were processed on the ground to validate the model-based pose estimation algorithms [11]. While the flight experiments are limited, there are various algorithms are proposed and validated by simulation in the literature. Some authors proposed using a point cloud obtained from range sensors to simultaneously estimate the kinematic state, geometry, and mass information of an unknown target [12]. In [13], a method that uses incremental smoothing and mapping system is proposed for mapping and estimation of the pose, the center of gravity (CG), and the inertia properties of an unknown uncooperative space target. The problem of CG and inertia properties' estimation and mapping was investigated in [14] and [15]. Others presented a real-time algorithm for pose estimation based on monocular Structure from Motion (SfM) [16], where a Bayesian filter is adopted to estimate the relative rotational dynamics. [16] makes a simplification that assumes a constant rotational velocity such that the target inertia matrix does not need to be estimated. In our prior work, we investigated the problem of feature extraction [17] and matching [18] for spacecraft pose estimation and presented Simultaneous Estimation of Pose and Shape (SEPS), which estimates the unknown space objects pose and shape using monocular vision and lidar [19]. SEPS features a sequential filter with attitude dynamics propagated using optical flow and is meant for a single chaser. A review in [9] compares and summarizes some of the existing relative pose estimation techniques using electro-optical sensors.

It is also relevant to note that the vision-based pose estimation algorithms for unknown space objects [20], [13] are also often compared with Simultaneous Localization and Mapping

(SLAM) approaches from terrestrial robotic platforms [21]. However, there are enough distinctions such that they require specialized algorithms. First, SLAM aims to estimate camera trajectories in a global frame, while the vision-based pose estimation algorithm estimates the relative motion of a moving target. Because the chaser poses may be assumed to be known, their trajectories are not part of the estimation problem. Second, pose estimation in space environments enjoys the benefit of well-characterized relative orbital mechanics and attitude dynamics compared to some other robot motion models in terrestrial applications. On the other hand, attitude estimation of an unknown space target has unique challenges associated with an unknown inertia matrix of the target, which may be unobservable depending on the rotational motion. This issue has to be addressed specifically. Lastly, the visual conditions experienced in space are much harsher compared to common terrestrial SLAM problems, resulting in high contrast, reflections, and time-varying shadows. For instance, Figures 2(a)-2(c) show the realistic rendering of an example target spacecraft (Cygnus) in Geostationary orbit (GEO) observed from three chaser spacecraft placed in a Passive Relative Orbit (Figure 2(d)). The images were created using Blender Cycle Engine [22] by ray-tracing a parallel ray from the Sun illuminating the 3D model of the target spacecraft (Cygnus) with textured surfaces. Figure 2 shows the effect of relative geometries between the target, chaser, and the sun. Some spacecraft experience worse visibility conditions than others, and the illumination varies over each orbit or each rotation of the target. Therefore relying on a single chaser may take a long time, if at all, before observing the full coverage.



**Figure 2.** Realistic rendering of a target spacecraft seen by three spacecraft in passive relative orbit (d).

Most of the existing work on vision-based pose estimation of an unknown target in the literature relies on the images collected by a single chaser. In this paper, we propose to overcome the known challenges by extending the single-spacecraft SEPS algorithm [19] to multi-spacecraft architecture, the M-SEPS. A multi-agent approach has multiple advantages. First, the chasers produce larger coverage of the target when they are placed in well spatially distributed relative orbits. The group of spacecraft can maintain a formation such that a persistent estimate can be produced. In this way, a virtual distributed camera system is created.

This is of special interest in those cases where some of the chasers cannot have good visual conditions but receive reasonable updates from the other chasers. The observations from different perspectives also improve the depth estimation, convergence rate, and accuracy. Moreover, if one of the chasers fails during the mission, the inspection task can still be completed by the other chasers. We present a cooperative strategy that generalizes to an arbitrary number of the chaser spacecraft.

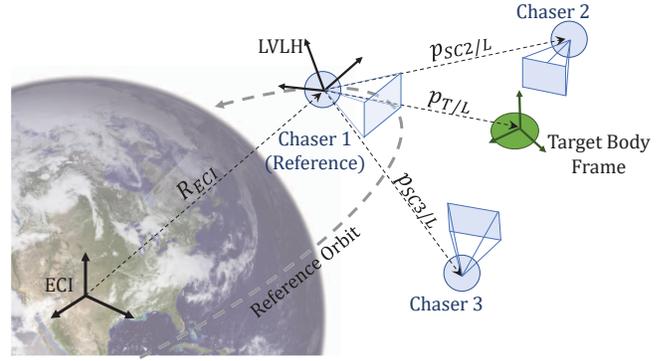
The goal of M-SEPS is to estimate both the attitude and CG of the target object using vision sensors on multiple chaser spacecraft. In particular, we tackle the problem in the distributed sensor network paradigm where the team of chaser spacecraft can exchange information over the local inter-spacecraft communication links. We exploit the recent development in distributed estimation theory [23], [24] where the approximate solution of a minimum variance estimate given the global information is computed in a distributed, iterative fashion. We highlight that the framework for a distributed sensor network is distinct from that of the cooperative SLAM in terrestrial applications, which assumes infrequent, event-based information exchange upon rendezvous. In the distributed sensor network paradigm, the independent measurements from decentralized platforms can be shared over the local communication network to be fused in an optimal fashion. As a related work [25], we recently presented the Decentralized Pose Estimation (DPE) and the Swarm Reference Frame Estimation (SRFE) algorithms, new approaches to decentralize the relative estimation using a spacecraft formation. The DPE improves the local formation estimate by considering the joint estimation over a relative sensing network, while the SRFE uses the information consensus approach [23], [24] to estimate the target position in a decentralized fashion. Both SRFE uses a Kalman information filter design, which poses clear advantages for multi-agent robotics perception, thus we adopt the same strategy in this work.

**Contributions:** The contributions of this paper are as follows: a) We present the multi-spacecraft algorithm architecture for the cooperative vision-based pose estimation of the uncooperative and unknown target, b) we exploit the theoretical tools from the distributed estimation of sensor networks to develop the extended decentralized information filter for the M-SEPS, we propose an improvement in dynamics update of the information matrix, leveraging the special structure that arises in the M-SEPS problem, and d) we validate the algorithm architecture through numerical simulations of relative orbits, measurements, and inter-spacecraft communications.

The rest of the paper is organized as follows. The estimation problem is formally stated in Section 2. Section 3 discusses the overall architecture of the M-SEPS, which includes key-point extraction, matching, and optical flow components. Section 4 discusses the decentralized information filter algorithm, which is a back-end filter that fuses distributed sensing information. Section 5 presents the validation of the approach through a simulation setup. Finally, conclusions are drawn in Section 6.

## 2. PROBLEM STATEMENT

A team of  $N$  spacecraft is orbiting in the proximity of a target in space. Each chaser  $i \in \mathcal{V}$  estimates the pose states of the target. The chasers are assumed to use a local communication network to exchange information, where the network topology is described by a simple undirected graph



**Figure 3.** Local-Vertical Local-Horizontal (LVLH) frame, defined on the reference spacecraft, and target's and chasers' body coordinate frames.

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}).$$

Figure 3 illustrates the conventions for the reference frames used in this paper. The absolute orbital motion of each chaser and target is described with respect to the Earth-Centered Inertial (ECI) frame. One of the chasers is chosen as a reference spacecraft, which defines a local-vertical local-horizontal (LVLH) coordinate frame. The target body frame is defined such that its origin coincides with its center of gravity (CG), which is to be estimated by M-SEPS. The orientation of the target frame is arbitrarily selected during the initialization since the target is unknown, and the attitude trajectory is described in terms of the relative attitude.

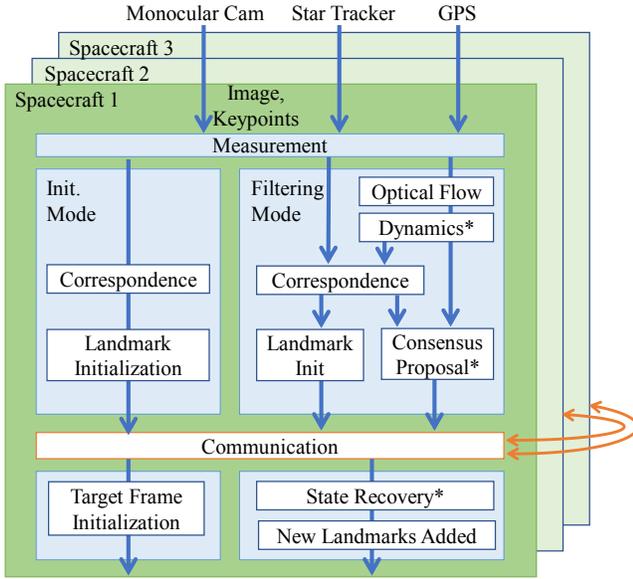
The state vector to be estimated in M-SEPS is  $\mathbf{x} = [\mathbf{x}_T; \mathbf{l}^1; \dots; \mathbf{l}^N]$  where  $\mathbf{x}_T$  is the target state and  $\mathbf{l}^i$  is the landmark states for  $i$ -th chaser. The target state  $\mathbf{x}_T = [\mathbf{p}_{T/L}; \mathbf{v}_{T/L}; \mathbf{q}_{T/I}]$  includes the relative positions and velocities of the target expressed in the LVLH frame, and its attitude in ECI frame. The angular velocity of the target is not a part of the state vector. Instead, they are obtained from the optical flow module explained in Section 3. The  $i$ -th chaser's landmark state is defined as  $\mathbf{l}^i = [\mathbf{p}_{1/T}^i; \dots; \mathbf{p}_{n_i/T}^i]$  where  $\mathbf{p}_{j/T}^i$  is  $j$ -th landmark visible to the  $i$ -th chaser, expressed in the target body frame. The resulting nonlinear chaser system is written as

$$\mathbf{x}_{T,k} = \mathbf{f}_T(\mathbf{x}_{T,k-1}, k) + \mathbf{w}_{T,k}, \quad (1)$$

$$\mathbf{p}_{j/T,k}^i = \mathbf{p}_{j/T,k-1}^i, \quad \forall i \in \mathcal{V}, j \in \{1, \dots, n^i\}, \quad (2)$$

$$\mathbf{z}_k^i = \mathbf{h}^i(\mathbf{x}_{T,k}, \mathbf{l}_k^i) + \mathbf{v}_k^i, \quad \forall i \in \mathcal{V}. \quad (3)$$

The nonlinear discrete-time target dynamics model  $\mathbf{f}_T$  may include any relevant environmental forces and control actuation. In this paper, we assume the coupling between the relative orbital mechanics and the attitude dynamics are negligible. The relative orbital propagation model includes the fully nonlinear equations for the Earth's point gravity and we assume torque-free attitude dynamics for the attitude propagation. The measurement  $\mathbf{z}_k^i$  are pixel coordinates of keypoints extracted from the image taken by  $i$ -th spacecraft at  $k$  using the computer vision algorithms. The measurement model  $\mathbf{h}^i$  projects the 3D landmarks to image plane and it is a function of the chaser pose (even though they are not included as the arguments to the function in Eq. 3 because these parameters are assumed to be known). The standard pinhole camera model is used for the camera projection model throughout this paper.  $\mathbf{w}_{T,k} \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_{T,k})$  and  $\mathbf{v}_k^i \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_k^i)$  for  $\forall i \in \mathcal{V}$



**Figure 4.** The collaborative pose estimation algorithm architecture. Each chaser has its own copy of the decentralized algorithm.

are process noise for target dynamics and measurement noise for each keypoint measurements, respectively.

Eqs. 1-3 can be equivalently written as a general nonlinear dynamical system, in terms of the full state vector  $\mathbf{x}$ , as follows

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, k-1) + \mathbf{w}_k, \quad (4)$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, k) + \mathbf{v}_k. \quad (5)$$

where  $\mathbf{w}_k = [\mathbf{w}_{T k}; \mathbf{0}; \dots; \mathbf{0}]$  and  $\mathbf{v}_k = [\mathbf{v}_k^1; \dots; \mathbf{v}_k^N]$  are the augmented noise vectors. If this process was to be estimated in a centralized algorithm, one can design a straightforward nonlinear observer such as EKF. The challenge of the distributed system is that each measurement  $\mathbf{z}^i$  is only available on  $i$ -th chaser. The goal of the M-SEPS is to approximate the minimum variance posterior estimate  $\mathbf{x}_k^+$  using the local communication network in a distributed fashion.

To simplify the formulation, we make the following assumptions. Each chaser is equipped with a star tracker and a Global Navigation Satellite System (GNSS) receiver [26] and that all chaser poses relative to the reference spacecraft are known. In order to have access to the real-time position of the reference spacecraft position, we assume an external estimation approach specifically designed for formation tracking. Next, we assume that the target is not applying relative translational maneuvers or such maneuver is negligible compared to the assumed process noise. Finally, we assume the chaser spacecraft have a connected but possibly time-varying communication graph.

### 3. M-SEPS ARCHITECTURE OVERVIEW

The Multi-spacecraft Simultaneous Estimation of Pose and Shape (M-SEPS) consists of multiple algorithm modules as shown in Figure 4. This section describes each of these modules, except the decentralized back-end filter, which is discussed in more detail in Section 4.

The M-SEPS can be primarily separated into *Initialization Mode* and *Filtering Mode* which are shown as two columns in Figure 4. We assume that the target CG and geometry are unknown or only partially known prior to the Initialization Mode. Therefore, Initialization Mode calculates both the initial set of landmarks and the target pose from a small batch of images. Also, the target body coordinate frame is initialized such that its origin is coincident with the estimated CG and it has an arbitrary orientation fixed with the target. During Filtering Mode, the state at the new time step is predicted by propagating the target dynamics and comparing it against the new observations. The Filtering Mode will continuously provide the relative pose estimate of the target such that a guidance and control module can use the product to proximity operation maneuvers in real-time. The modules with asterisks in Figure 4 are part of the extended decentralized information filter described in Section 4.

The algorithms in each Initialization and Filtering Modes are also split between the *pre-communication*, *communication*, and *post-communication* steps, where the communication aggregates information shared across multiple spacecraft. The following sections visit each module and discuss its functionality.

#### Keypoint Extraction and Correspondence

The strategies for keypoint correspondences are different for Initialization and Filtering Modes. During the Initialization Mode, the two sets of extracted keypoints from two different epochs are matched by all-to-all, brute force matching, followed by ratio-test. In addition, we apply a random sample consensus (RANSAC) algorithm with a 5-point Stewenius algorithm to compute the 3D landmark position, which is discussed in Section 3. During the Filtering Mode, the extracted keypoints are matched with projected landmarks. The search region is reduced by using the predicted keypoint location and covariance of the landmarks. We assume the standard pinhole camera model to describe the projection of the 3D landmarks onto the image plane.

Each chaser searches for correspondences only among images taken by itself and not across the multiple spacecraft. This architectural choice is motivated by a few reasons. First, finding correspondences of descriptors across multiple spacecraft requires communicating the set of descriptors, which increases the bandwidth. Second, even with feasible communication, the chance of detecting a correspondence is small across multiple spacecraft, given the large variation in the visual conditions when the chasers are spatially well distributed.

#### Target Frame Initialization

The primary function of the Initialization Mode is to initialize the target frame, which involves selecting the origin and orientation of the frame consistent among all chasers and defining all the detected inliers in the target frame. The target frame initialization involves three steps: individual initialization, communication, and transformation into a common frame.

First, individual spacecraft initialize the respective landmarks expressed in its camera frame prior to the communication step. Given the camera pose in ECI at each epoch is known from GNSS and the star tracker, the target pose trajectory is solved along with their landmarks. To solve this, the 5-point Stewenius algorithm [27] with random sample consensus (RANSAC) is used to determine inliers, followed by least-

square optimization. With relative pose transformation from the chaser’s camera to the reference’s camera, the landmark positions are initialized in the reference spacecraft frame. At this time, the reference spacecraft also initializes the target frame by selecting an arbitrary attitude and a coarse estimate of the target CG. The coarse estimate of the target CG is defined as one of the landmark features observed.

Second, spacecraft exchange information in the communication step. The landmarks estimated from each chaser and the target initial frame from the reference spacecraft are shared with all the spacecraft. Finally, each spacecraft applies the pose transformation to obtain landmarks in the target reference frame.

#### Optical Flow

The optical flow module uses the sequence of images to compute the angular velocity in a similar way as done in the previous work [19]. When the initial frame and the inertia matrix of the target are unknown, direct observation of angular measurement has advantages such as simpler propagation of attitude quaternion and avoiding estimation of the inertia matrix. The optical flow may be obtained by classical methods such as Lucas-Kanade [28] or more recent methods such as using a convolutional neural network trained with sequences of images [29].

#### Communication

The information shared among neighbor spacecraft is different for Initialization and Filtering Modes. The communication module manages the information exchange between neighboring spacecraft. Given the edges in the communication graph  $\mathcal{E}^k$ , the communication link is established between  $(i, j) \in \mathcal{E}^k$ . We assume that the communication rate is higher than the estimation rate, i.e. there may be multiple communication exchanges at each filter time epoch.

#### Landmark Initialization

If there is a new keypoint correspondence that was not previously tracked, the landmark state is added to the filter. First, individual spacecraft detects new landmarks prior to the communication step. The new landmarks’ 3D positions are broadcast to all the spacecraft, and finally, all the new landmarks from the epoch are added to the state. This step is implemented in a similar way to SEPS [19].

## 4. EXTENDED DECENTRALIZED INFORMATION FILTER

Section 3 provided an overview of sub-modules that enable the computer vision pipeline for M-SEPS. This section discusses the back-end filter extended decentralized information filter (EDIF), a nonlinear extension of the decentralized information filter.

Should there be a centralized node that has access to all the measurements in the network, a standard extended information filter in Algorithm 1 is sufficient to fuse measurements from multiple sensors. In Algorithm 1,  $\mathbf{j}_k$  and  $\mathbf{J}_k$  denote the information vector and information matrix at time  $k$  and  $\mathbf{F}_k$  and  $\mathbf{H}_k$  are Jacobians of dynamics and measurement functions.  $\mathbf{W}_k$  and  $\mathbf{V}_k$  are covariance matrices for the process noise and the measurement noise, respectively. The superscript “+” and “-” denotes posterior and prior estimates. Notice that the information filter form of the Kalman

---

### Algorithm 1: Extended Information Filter

---

**Result:**  $\mathbf{x}_k^+, \mathbf{J}_k^+$

1 Prediction step

$$\mathbf{x}_k^- = \mathbf{f}(\mathbf{x}_{k-1}^+, k), \quad (6)$$

$$(\mathbf{J}_k^-)^{-1} = \mathbf{F}_k (\mathbf{J}_{k-1}^+)^{-1} \mathbf{F}_k^\top + \mathbf{W}_k, \quad (7)$$

2 Compute the information vector

$$\mathbf{j}_k^- = \mathbf{J}_k^- \mathbf{x}_k^- \quad (8)$$

3 Get measurement  $\mathbf{z}_k$

4 Measurement update

$$\mathbf{j}_k^+ = \mathbf{j}_k^- + \mathbf{H}_k^\top \mathbf{V}_k^{-1} (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^-, k) + \mathbf{H}_k \mathbf{x}_k^-), \quad (9)$$

$$\mathbf{J}_k^+ = \mathbf{J}_k^- + \mathbf{H}_k^\top \mathbf{V}_k^{-1} \mathbf{H}_k \quad (10)$$

5 Recover the state vector

$$\mathbf{x}_k^+ = (\mathbf{J}_k^+)^{-1} \mathbf{j}_k^+ \quad (11)$$


---

filter admits a simple summation form for the measurement equations (Eqs. 9 and 10) facilitating a distributed implementation.

Algorithm 1 in the current form has some disadvantages, however. First, the algorithm is still not distributed. Second, the matrix inverse operation of the information matrix is computationally intensive as the state vector size becomes large. This is a particular concern for M-SEPS where the state includes landmark states. We will exploit the Decentralized Information Filter [24] which obtains the approximate solution to the minimum variance estimate for a linear system. EDIF will extend this result to nonlinear dynamics and measurement models. The following sections discuss how to distribute the measurement update. We also develop an improvement to the dynamics update of the information matrix which leverages the special structure of the M-SEPS problem.

#### Measurement Update and Consensus

When the overall measurement  $\mathbf{z}_k$  consists of independent measurements, the measurement update becomes a simple sum of the measurement contributions from all the observations.

$$\mathbf{j}_k^+ = \mathbf{j}_k^- + \sum_{i \in \mathcal{V}} \mathbf{H}_k^i \mathbf{V}_k^{i-1} (\mathbf{z}^i - \mathbf{h}^i(\mathbf{x}_k^-, k) + \mathbf{H}_k^i \mathbf{x}_k^-), \quad (12)$$

$$\mathbf{J}_k^+ = \mathbf{J}_k^- + \sum_{i \in \mathcal{V}} \mathbf{H}_k^i \mathbf{V}_k^{i-1} \mathbf{H}_k^i, \quad (13)$$

where  $\mathbf{z}^i$  and  $\mathbf{h}^i$  denote the measurement and measurement model of the keypoints as seen by the  $i$ -th chaser. The information update has a block-sparse structure that allows further simplifications. Recall that the state vector is divided into sub-blocks  $\mathbf{x} = [\mathbf{x}_T; \mathbf{1}^1; \dots; \mathbf{1}^N]$ . After sub-dividing the information vector and matrix into corresponding block elements, we can equivalently write Eqs. 12 and 13 as following. For simplicity, we drop the subscript  $k$  which denoted the

time.

$$\mathbf{J}_{TT}^+ = \mathbf{J}_{TT}^- + \sum_{i \in \mathcal{V}} \mathbf{H}_{iT}^\top \mathbf{V}_k^{i-1} \mathbf{H}_{iT}, \quad (14)$$

$$\mathbf{J}_{ii}^+ = \mathbf{J}_{ii}^- + \mathbf{H}_{ii}^\top \mathbf{V}_k^{i-1} \mathbf{H}_{ii}, \quad \forall i \in \mathcal{V}, \quad (15)$$

$$\mathbf{J}_{iT}^+ = \mathbf{J}_{iT}^- + \mathbf{H}_{ii}^\top \mathbf{V}_k^{i-1} \mathbf{H}_{iT}, \quad \forall i \in \mathcal{V}, \quad (16)$$

$$\mathbf{J}_{ij}^+ = \mathbf{J}_{ij}^-, \quad \forall i, j \in \mathcal{V}, i \neq j. \quad (17)$$

The linear form of the measurement update in the information filter leads to the distributed implementation. For a formation with a small number of chasers, Eqs. 12-13 may be implemented by simply relaying the contributions from individual terms. For a formation with a large number of spacecraft, a consensus algorithm may be used to iteratively converge to average. To see this, we manipulate the measurement update equation to

$$\mathbf{j}_k^+ = \frac{1}{N} \sum_{i \in \mathcal{V}} \left( \mathbf{j}_k^- + N \mathbf{H}_k^{i \top} \mathbf{V}_k^{i-1} (\mathbf{z}^i - \mathbf{h}^i(\mathbf{x}_k^-, k) + \mathbf{H}_k^i \mathbf{x}) \right), \quad (18)$$

$$\mathbf{J}_k^+ = \frac{1}{N} \sum_{i \in \mathcal{V}} \left( \mathbf{J}_k^- + N \mathbf{H}_k^{i \top} \mathbf{V}_k^{i-1} \mathbf{H}_k^i \right), \quad (19)$$

where  $N = \text{card}(\mathcal{V})$ . Assuming an undirected graph for the communication topology, the decentralized information filter applies the following consensus protocol to compute the above equations.

$$\mathbf{u}_{c+1}^i = \mathbf{u}_c^i + \epsilon \sum_{j \in \mathcal{N}(i)} (\mathbf{u}_c^j - \mathbf{u}_c^i), \quad (20)$$

$$\mathbf{U}_{c+1}^i = \mathbf{U}_c^i + \epsilon \sum_{j \in \mathcal{N}(i)} (\mathbf{U}_c^j - \mathbf{U}_c^i), \quad (21)$$

where  $\epsilon$  is the constant design parameter called the consensus coefficient. Later, Theorem 1 shows that  $\epsilon$  must be sufficiently small to guarantee convergence. At the first iteration of consensus, the  $\mathbf{u}_c^i$  and  $\mathbf{U}_c^i$  are initialized as

$$\mathbf{u}_0^i = \mathbf{j}_k^- + N \mathbf{H}_k^{i \top} \mathbf{V}_k^{i-1} (\mathbf{z}^i - \mathbf{h}^i(\mathbf{x}_k^-, k) + \mathbf{H}_k^i \mathbf{x}), \quad (22)$$

$$\mathbf{U}_0^i = \mathbf{J}_k^- + N \mathbf{H}_k^{i \top} \mathbf{V}_k^{i-1} \mathbf{H}_k^i. \quad (23)$$

The following theorem guarantees the convergence of the consensus protocol given that  $\epsilon$  is bounded by a function of the degree of the graph.

**Theorem 1** (Convergence of consensus protocol [30]) Consider a network of agents with communication graph  $\mathcal{G}$  applying the consensus algorithms Eqs. 20-21 where  $0 < \epsilon < \frac{1}{\Delta}$  and  $\Delta$  is the maximum degree of the network. If the digraph is balanced, an average-consensus is asymptotically reached. That is  $\lim_{c \rightarrow \infty} \mathbf{u}_c^i = \frac{1}{N} \sum_{j \in \mathcal{V}} \mathbf{u}_0^j$  for  $\forall i \in \mathcal{V}$ .

The maximum degree of the network may be controlled by specifying the maximum degree for each node, such that  $\epsilon$  may be selected prior to the mission. Also, in practice, such a consensus algorithm has been shown to have sufficient convergence with finite iterations [25].

### Information Time-Update

One drawback of using the information filter as opposed to the Kalman filter is the time-update of the information matrix

requires the additional inversion of the covariance matrix. The computational complexity of inverse scales with  $O(n^3)$  in a fully dense matrix where  $n$  is the dimension of the state vector. For a large scale problem like M-SEPS where  $n$  is large, reducing the inversion operation is desirable. With this in mind, we introduce the following proposition in which the structure of the M-SEPS problem is used to reduce the complexity of computation.

**Proposition 1:** Assume  $\mathbf{J}_{k-1}^+$  and  $\mathbf{W}_{T_k}$  are positive definite matrices. Suppose  $\mathbf{F}_{T_k}$  and  $\mathbf{F}_k$  corresponds to Jacobians of  $\mathbf{f}_T$  and  $\mathbf{f}$  respectively, and  $\mathbf{W}_k$  is the process noise covariance at  $k$ . Then, the time-update step of the information matrix can be re-written as

$$\mathbf{J}_k^- = \bar{\mathbf{J}} - \bar{\mathbf{J}} \boldsymbol{\omega} (\mathbf{I} + \boldsymbol{\omega}^\top \bar{\mathbf{J}} \boldsymbol{\omega})^{-1} \boldsymbol{\omega}^\top \bar{\mathbf{J}} \quad (24)$$

where

$$\bar{\mathbf{J}} := \mathbf{F}_k^{-\top} \mathbf{J}_{k-1}^+ \mathbf{F}_k^{-1}, \quad (25)$$

$$\boldsymbol{\omega} := [(\mathbf{W}_{T_k})^{\frac{1}{2}}; \mathbf{0}; \dots; \mathbf{0}]. \quad (26)$$

*Proof:* The most general time-update equation of the information matrix is given by

$$(\mathbf{J}_k^-)^{-1} = \mathbf{F}_k (\mathbf{J}_{k-1}^+)^{-1} \mathbf{F}_k^\top + \mathbf{W}_k \quad (27)$$

where  $\mathbf{F}$  is the Jacobian of the overall dynamics  $\mathbf{f}(\mathbf{x})$ . We leverage that the process noise is introduced only to the target state  $\mathbf{x}_T$  and not on the landmarks. Suppose  $\mathbf{W}_T$  is the process noise covariance corresponding to the target propagation. Then, the process noise covariance for overall dynamics may be written as  $\mathbf{W} = \boldsymbol{\omega} \boldsymbol{\omega}^\top$  where  $\boldsymbol{\omega}$  is defined above with zero matrices, appropriately sized. Let  $\bar{\mathbf{J}} = \mathbf{F}^{-\top} \mathbf{J}_{k-1}^+ \mathbf{F}^{-1}$ , where dynamics of Jacobian has block diagonal structure  $\mathbf{F} = \text{diag}(\mathbf{F}_T, \mathbf{I}, \dots, \mathbf{I})$ . Using the definitions of  $\boldsymbol{\omega}$  and  $\bar{\mathbf{J}}$ , Eq. 27 may be equivalently written as

$$\begin{aligned} \mathbf{J}_k^- &= (\bar{\mathbf{J}}^{-1} + \boldsymbol{\omega} \boldsymbol{\omega}^\top)^{-1} \\ &= \bar{\mathbf{J}} - \bar{\mathbf{J}} \boldsymbol{\omega} (\mathbf{I} + \boldsymbol{\omega}^\top \bar{\mathbf{J}} \boldsymbol{\omega})^{-1} \boldsymbol{\omega}^\top \bar{\mathbf{J}}. \end{aligned} \quad (28)$$

The second equality holds by using the matrix inversion lemma. ■

We note that the computational complexity of Eq. 24 with respect to landmark size scales better than Eq. 27. Jacobian  $\mathbf{F}$  has a block diagonal structure with mostly identity elements, so the inverse of Jacobian is simply  $\mathbf{F}^{-1} = \text{diag}(\mathbf{F}_T^{-1}; \mathbf{I}; \dots; \mathbf{I})$ . The computation of  $\boldsymbol{\omega}$  does not depend  $n$ , and computation  $\bar{\mathbf{J}}$  and  $\bar{\mathbf{J}} \boldsymbol{\omega}$  both scales linearly, i.e.  $O(n)$ . The matrix inversion  $(\mathbf{I} + \boldsymbol{\omega}^\top \bar{\mathbf{J}} \boldsymbol{\omega})^{-1}$  is order of  $n_T \times n_T$  matrix and does not depend on  $n$ . The operation that requires most computation is  $\boldsymbol{\omega} (\mathbf{I} + \boldsymbol{\omega}^\top \bar{\mathbf{J}} \boldsymbol{\omega})^{-1} \boldsymbol{\omega}^\top \bar{\mathbf{J}}$  which requires  $O(n^2)$  computation. Therefore this manipulation eliminates the inversion of the full information matrix.

### State Recovery

Once the information vector and the information matrix are computed, we need to compute the state vector, which involves the inversion of the information matrix

$$\mathbf{x} = (\mathbf{J}^+)^{-1} \mathbf{j}^+. \quad (29)$$

---

**Algorithm 2:** Decentralized Extended Information Filter

---

**Result:**  $\mathbf{x}_k^+$ ,  $\mathbf{J}_k^+$ 

1 Prediction step

$$\mathbf{x}_{T_k}^- = \mathbf{f}_T(\mathbf{x}_{T_{k-1}}^+), \quad \mathbf{I}_{i,k}^- = \mathbf{I}_{i,k-1}^+ \quad (30)$$

$$\bar{\mathbf{J}} = \mathbf{F}^{-\top} \mathbf{J}_{k-1}^+ \mathbf{F}^{-1} \quad (31)$$

$$\mathbf{J}_k^- = \bar{\mathbf{J}} - \bar{\mathbf{J}} \boldsymbol{\omega} (\mathbf{I} + \boldsymbol{\omega}^\top \bar{\mathbf{J}} \boldsymbol{\omega})^{-1} \boldsymbol{\omega}^\top \bar{\mathbf{J}} \quad (32)$$

2 Get measurement  $\mathbf{z}_i$ 

3 Compute consensus proposals

$$\mathbf{u}_0^i = \mathbf{j}^{i-} + N \mathbf{H}_k^{i\top} \mathbf{V}_i^{-1} (\mathbf{z}^i - \mathbf{h}^i(\mathbf{x}_k^-, k) + \mathbf{H}_k^i \mathbf{x}_k^-) \quad (33)$$

$$\mathbf{U}_0^i = \mathbf{J}^{i-} + N \mathbf{H}_k^{i\top} \mathbf{V}_i^{-1} \mathbf{H}_k^i \quad (34)$$

**while**  $c \leq C$  **do**

4 Perform consensus

$$\mathbf{u}_c^i = \mathbf{u}_{c-1}^i + \epsilon \sum_{j \in \mathcal{N}_i} (\mathbf{u}_{c-1}^j - \mathbf{u}_{c-1}^i) \quad (35)$$

$$\mathbf{U}_c^i = \mathbf{U}_{c-1}^i + \epsilon \sum_{j \in \mathcal{N}_i} (\mathbf{U}_{c-1}^j - \mathbf{U}_{c-1}^i) \quad (36)$$

**send**

6 Compute the posterior state and information matrix

$$\mathbf{x}_k^+ = (\mathbf{U}_C^i)^{-1} \mathbf{u}_C^i \quad (37)$$

$$\mathbf{J}^+ = \mathbf{U}_C^i \quad (38)$$

7 Add new landmark states and covariances

---

*Tracked Landmarks*

The newly discovered landmark states and respective covariances are initialized after the measurement update step. This takes place after the communication step, ensuring that new landmarks detected by any of the chasers are included in the new state vector and the information matrix. The information matrix is updated by placing the inverse of the initial covariance of the new landmarks on the extended diagonal block.

This completes the discussion of all the necessary components of the EDIF algorithm. EDIF algorithm is summarized in Algorithm 2 for clarity. In context of the overall architecture shown in Figure 4, the modules that are primarily involved with the EDIF are denoted with asterisks.

## 5. SIMULATION

The part of the M-SEPS architecture described in Section 3-4 is validated in a computer simulation. The purpose of the simulation is to validate the multi-spacecraft architecture, the distributed algorithm, and the dynamic allocation of landmark states and covariances. As such, we do not use the images to extract the keypoints and synthetic images are only used for visualization purposes. We assume that the keypoints extraction, keypoint correspondence, and optical flow are solved by functional sub-modules. We refer the

Orbital Parameter	Value
Semi-major axis	42167.0 km
Eccentricity	2.9E-4
Inclination	8.1E-2 deg
Argument of perigee	354.2 deg
RAAN	68.8 deg
True anomaly	240.8 deg

**Table 1.** Geostationary orbital parameters

Relative State	Chaser 1	Chaser 2	Chaser 3
Position [m]	20.0	-10.0	-10.0
	0.0	-34.6	34.6
Velocity [mm/s]	34.6	-17.3	-17.3
	0.0	-1.26	1.26
Velocity [mm/s]	-2.92	-1.46	1.46
	0.0	-2.19	2.19

**Table 2.** Chaser initial states in target-LVLH frame

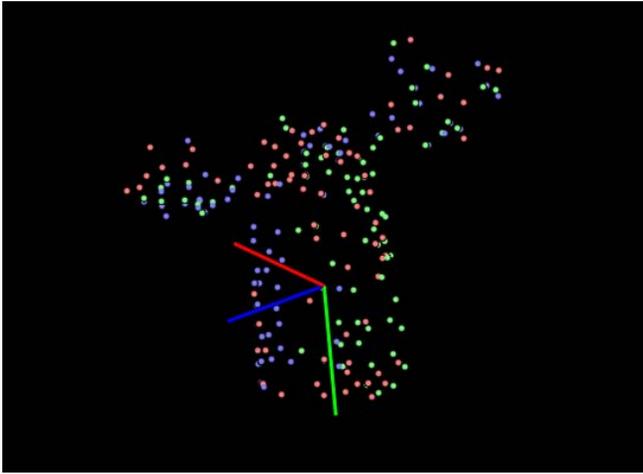
reader to [19], [29] for further implementation details. Future work includes validation of the algorithm with the computer vision algorithm in the loop using the realistic synthetic images.

Extracted keypoints in this simulation are obtained by projecting a set of pre-defined 3D landmarks attached to the exterior of the spacecraft model to each chaser's image plane. A set of keypoints was simulated for each chaser based on the relative pose between the chaser and the target at the time. Only the landmarks that have line-of-sight are considered visible. The keypoints corresponding to the same landmark have the same descriptors at different epoch, but the descriptors are different across different spacecraft. At each epoch, each chaser processes the observation and runs an iteration of the filter. The swarm of spacecraft only shared the variables over the simulated inter-spacecraft communication. The algorithms on all the chaser spacecraft are assumed to run in a synchronized fashion via clock synchronization obtained from the GNSS. The differential GNSS and the star tracker measurements used in the formation flight estimation are simulated by the ground truth relative state.

*Orbital Mechanics and Attitude Dynamics*

The target and chaser spacecraft are placed in geostationary orbits and the target's initial orbital parameters are tabulated in Table 1. The initial positions and velocities of the chaser spacecraft were selected such that they are in a formation with respect to the target and each other, as shown in Table 2. The resulting trajectories of the chasers form concentric circular orbits with respect to the target as shown in Figure 2(d). The absolute orbit for each spacecraft was computed by propagating the respective spacecraft in the Earth-Centered Inertial (ECI) frame. The target spacecraft simulated for this validation has a slight tumble which makes the CG position in spin-parallel direction still observable over a longer period of time.

The target's initial angular velocity is 1 degree per second. While the target attitude is numerically propagated by torque-free dynamics, each chaser is assumed to track a smooth attitude trajectory that satisfies its pointing requirement. Specifically, each chaser points its camera optical axis towards the target spacecraft such that the target is always entirely within



**Figure 5.** Reconstructed shape obtained from the tracked landmarks. Red, blue, and green markers denote landmarks seen by Chasers 1, 2, and 3, respectively.

the field of view. We assume that reaction wheels will provide the necessary slew maneuvers. Note that because chaser spacecraft inertial poses are assumed to be known from the GNSS and the star tracker, the M-SEPS does not need to model the chasers' dynamics on-board.

#### *Simulated Measurements*

Each chaser receives the GPS and star tracker measurements as the pose with respect to the ECI frame. These measurements are simulated as the ground truth positions. Note that this is not a strong assumption, as in GEO and above, although above the GPS constellation, it is possible to track GPS signals for navigation using high sensitivity receivers [31]. Millimeter accuracy has been proved in LEO filtering differential GPS observations with a model of the spacecraft dynamics. The same accuracy is theoretically also possible in higher orbits, such as in GEO, by tightly fusing the GPS observations with measurements of another sensor [26].

We also assume that the keypoint detection and correspondence are solved by the front-end computer vision algorithm, instead of extracting the keypoints from synthetic images using the computer vision algorithms.

A set of keypoints are simulated by projecting the landmark to the camera origin given the relative pose of the target with respect to the camera at each time epoch. A set of pre-defined 3D landmarks are placed on the exterior of the spacecraft model. A pinhole camera model is used and the landmarks are only visible when there is no obstruction on their line of sight. We artificially add Gaussian noise to each keypoint observation at each epoch. The image size is 1024-by-1280 pixels and the focal length is  $f = 2400$  pixels.

#### *Results*

The reconstructed shape of the target is shown as a 3D point cloud in Figure 5. Even though the target spacecraft (Cygnus) has a complex geometry including deployed solar panels, the reconstructed point cloud closely resembles the target shape. The red, blue, and green markers denote the landmarks that are visible by Chaser 1, 2, and 3 respectively at the epoch. Because three spacecraft are spatially distributed, they cover different surfaces of the Cygnus, showing the advantage of the multi-spacecraft approach. Figure 1 shows the same



**Figure 6.** Tracked landmark overlaid on the synthetic image. Circles around each landmark indicate the size of uncertainties. Green landmarks indicate states that are tracked for more than 10 consecutive frames.

reconstructed shape from another perspective. The figure also includes the camera pose trajectory expressed in the estimated target reference frame. Even though the relative trajectories of the chasers were designed to be concentric circular orbit in the LVLH frame (Figure 2(d)), the chaser trajectories are more complex when expressed in the target reference frame due to the target's own rotation.

The projected keypoints of Chaser 1 is shown along with the synthetic image in Figure 6. For each keypoint, the corresponding landmark covariance projected onto the image plane is shown as a circle around the keypoint. Green color indicates that the keypoint was tracked for 10 or more consecutive frames, while orange color indicates the point was tracked for less. The figure shows that the landmarks with longer tracks have smaller covariances as expected.

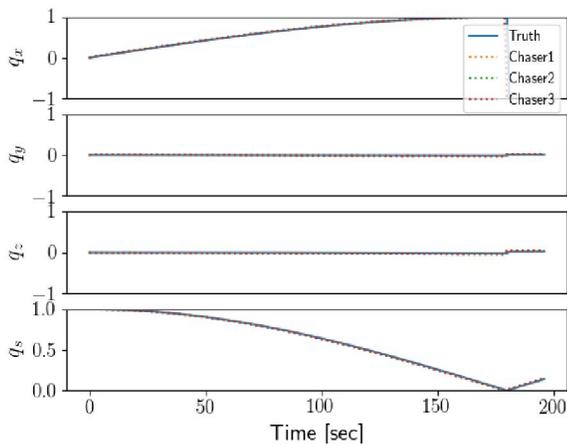
The results of target pose tracking are shown in Figures 7 and 8. Figure 7 shows the quaternion values of both truth and estimated target attitude, with respect to the initial frame. The estimated attitude quaternion follows closely of the true attitude. It also shows that the estimate obtained by all chasers agree with each other. Figure 8 shows the estimation error for target's CG. The error is defined with respect to the ground truth target CG which was used to generate the simulation. The CG location estimation errors are described in terms of the parallel and perpendicular directions to the target spin-axis. This projection is selected because the observability of the CG location in the perpendicular direction is expected to be higher than the direction parallel for any rotating object with constant or slowly varying angular velocity vector. Figure 8 verifies these behaviors where it shows a quick initial convergence in both parallel and perpendicular directions. Then the position estimate continues to converge (approach towards zero) with higher error in the parallel direction. The CG position estimate remains bounded over time.

## 6. CONCLUSIONS

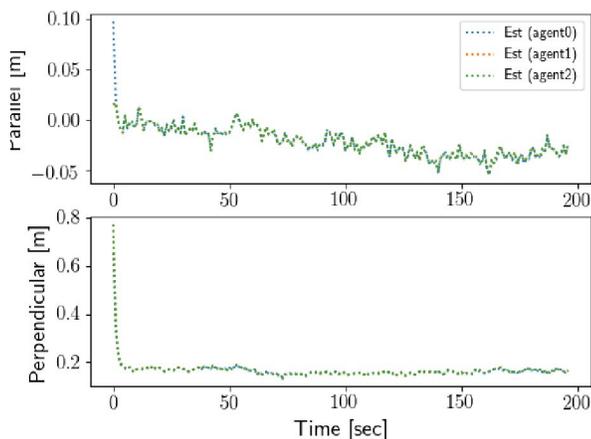
We presented the cooperative vision-based pose estimation algorithm called Multi-spacecraft Simultaneous Estimation of Pose and Shape (M-SEPS). M-SEPS is posed as a distributed sensor network paradigm. Using inter-spacecraft communication, M-SEPS tightly integrates the vision-based

## REFERENCES

- [1] J. L. Goodman, "History of space shuttle rendezvous and proximity operations," *Journal of Spacecraft and Rockets*, vol. 43, no. 5, pp. 944–959, 2006.
- [2] T. E. Rumford, "Demonstration of autonomous rendezvous technology (dart) project summary," in *Space Systems Technology and Operations*, vol. 5088. International Society for Optics and Photonics, 2003, pp. 10–19.
- [3] S. Van Winkle, "Advanced video guidance sensor (avgs) project summary," in *Spaceborne Sensors*, vol. 5418. International Society for Optics and Photonics, 2004, pp. 10–20.
- [4] R. B. Friend, "Orbital express program summary and mission overview," in *Sensors and Systems for space applications II*, vol. 6958. International Society for Optics and Photonics, 2008, p. 695803.
- [5] R. Pinson, R. Howard, and A. Heaton, "Orbital express advanced video guidance sensor: ground testing, flight results and comparisons," in *AIAA guidance, navigation and control conference and exhibit*, 2008, p. 7318.
- [6] L. Blarre, N. Perrimon, C. Moussu, P. Da Cunha, and S. Strandmoe, "Atv videometer qualification," in *55th International Astronautical Congress of the International Astronautical Federation, the International Academy of Astronautics, and the International Institute of Space Law*, 2004, pp. A–3.
- [7] G. Aglietti, B. Taylor, S. Fellowes, T. Salmon, A. Hall, T. Chabot, C. Bernal, and A. Pisseloup, "Remove debris mission: In-orbit operations," 2019.
- [8] S. Pedrotty, J. Sullivan, E. Gambone, and T. Kirven, "Seeker free-flying inspector gnc system overview," 02 2019.
- [9] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "A review of cooperative and uncooperative spacecraft pose determination techniques for close-proximity operations," *Progress in Aerospace Sciences*, vol. 93, pp. 53–72, 2017.
- [10] D. Fourie, B. E. Tweddle, S. Ulrich, and A. Saenz-Otero, "Flight results of vision-based navigation for autonomous spacecraft inspection of unknown objects," *Journal of Spacecraft and Rockets*, vol. 51, no. 6, pp. 2016–2026, 2014.
- [11] J. L. Jørgensen and M. Benn, "Vbs-the optical rendezvous and docking sensor for prisma," *NordicSpace*, pp. 16–19, 2010.
- [12] M. D. Lichter and S. Dubowsky, "State, shape, and parameter estimation of space objects from range images," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 3. IEEE, 2004, pp. 2974–2979.
- [13] B. E. Tweddle, A. Saenz-Otero, J. J. Leonard, and D. W. Miller, "Factor graph modeling of rigid-body dynamics for localization, mapping, and parameter estimation of a spinning object in space," *Journal of Field Robotics*, vol. 32, no. 6, pp. 897–933, 2015.
- [14] T. P. Setterfield, D. W. Miller, J. J. Leonard, and A. Saenz-Otero, "Mapping and determining the center of mass of a rotating object using a moving observer," *The International Journal of Robotics Research*, vol. 37, no. 1, pp. 83–103, 2018.
- [15] T. P. Setterfield, D. W. Miller, A. Saenz-Otero, E. Fraz-



**Figure 7.** Target attitude quaternion tracking result



**Figure 8.** Target center of gravity estimation error

feature tracking problem with a distributed estimation framework. We provided an overview of the multi-spacecraft algorithm architecture that consists of the computer vision pipeline, communication, and back-end filtering. We proposed the extended decentralized information filtering that approximately solves the minimum variance estimate of the global information but implemented in a distributed fashion, and proposed an improvement in computation that exploits the special structure of the M-SEPS problem. We validated the distributed algorithm and some of the algorithm pipelines using the simulation.

## ACKNOWLEDGMENTS

Kai Matsuka was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1745301. Vincenzo Capuano was supported by the Swiss National Science Foundation. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Copyright 2020 California Institute of Technology. U.S. Government sponsorship acknowledged.

- zoli, and J. J. Leonard, "Inertial properties estimation of a passive on-orbit object using polhode analysis," *Journal of Guidance, Control, and Dynamics*, vol. 41, no. 10, pp. 2214–2231, 2018.
- [16] S. Augenstein and S. M. Rock, "Improved frame-to-frame pose tracking during vision-only slam/sfm with a tumbling target," in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 3131–3138.
- [17] V. Capuano, S. R. Alimo, A. Q. Ho, and S.-J. Chung, "Robust features extraction for on-board monocular-based spacecraft pose acquisition," in *AIAA Scitech 2019 Forum*, 2019, p. 2005.
- [18] A. Harvard, V. Capuano, E. Y. Shao, and S.-J. Chung, "Spacecraft pose estimation from monocular images using neural network based keypoints and visibility maps."
- [19] V. Capuano, K. Kim, A. Harvard, and S.-J. Chung, "Monocular-based pose determination of uncooperative space objects," *Acta Astronautica*, vol. 166, pp. 493–506, 2020.
- [20] A. Sonnenburg, M. Tkocz, and K. Janschek, "EKF-slam based approach for spacecraft rendezvous navigation with unknown target spacecraft," *IFAC Proceedings Volumes*, vol. 43, no. 15, pp. 339–344, 2010.
- [21] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Transactions on robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.
- [22] B. O. Community, *Blender - a 3D modelling and rendering package*, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. [Online]. Available: <http://www.blender.org>
- [23] A. T. Kamal, J. A. Farrell, and A. K. Roy-Chowdhury, "Information weighted consensus filters and their application in distributed camera networks," *IEEE Transactions on Automatic Control*, vol. 58, no. 12, pp. 3112–3125, 2013.
- [24] S. Bandyopadhyay and S.-J. Chung, "Distributed bayesian filtering using logarithmic opinion pool for dynamic sensor networks," *Automatica*, vol. 97, pp. 7–17, 2018.
- [25] K. Matsuka, A. O. Feldman, S. Lupu, S.-J. Chung, and F. H. Hadaegh, "Decentralized formation pose estimation for spacecraft swarms," *Advances in Space Research*.
- [26] V. Capuano, A. Harvard, Y. Lin, and S.-J. Chung, "DGNSS-vision integration for robust and accurate relative spacecraft navigation," in *Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019)*, Miami, Florida, 2019, pp. 2923–2939.
- [27] H. Li and R. Hartley, "Five-point motion estimation made easy," in *18th International Conference on Pattern Recognition (ICPR'06)*, vol. 1. IEEE, 2006, pp. 630–633.
- [28] B. D. Lucas, T. Kanade, *et al.*, "An iterative image registration technique with an application to stereo vision," 1981.
- [29] S. Lee, V. Capuano, A. Harvard, and S.-J. Chung, "Fast uncertainty estimation for deep learning based optical flow," in *International Conference on Intelligent Robots and Systems (IROS) 2020, Las Vegas, NV, USA*, 2020.
- [30] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
- [31] V. Capuano, C. Botteron, J. Leclère, J. Tian, Y. Wang, and P.-A. Farine, "Feasibility study of GNSS as navigation system to reach the moon," *Acta Astronautica*, vol. 116, pp. 186–201, 2015.

## BIOGRAPHY



**Kai Matsuka** is a Ph.D. Candidate in the Graduate Aerospace Laboratories of the California Institute of Technology (GALCIT). He received his B.S. in Aerospace Engineering from the University of California, Los Angeles in 2016 and M.S. in Space Engineering from Caltech in 2018. Before attending Caltech, he was a Guidance and Control Analysis engineer at NASA Jet Propulsion Laboratory where he did research on fuel-efficient control of formation flying for the Earth science missions. His current research focuses on distributed estimation, multi-agent robotic systems, and formation flying spacecraft. He is a recipient of the National Science Foundation Graduate Research Fellowship.



**Angel Santamaria-Navarro** is a Robotics Research Technologist in the Robotic Aerial Mobility Group at NASA-JPL. He received the BSc degrees in T. Telecommunications and Industrial Management, and MSc and PhD in Automatic Control, Robotics and Computer Vision from the "Universitat Politècnica de Catalunya" (UPC), Barcelona, in 2004, 2012 and 2017 respectively. He was a finalist of the Georges Giralt PhD award for the best European thesis in robotics (2018) and he has published a book, several other book chapters and top international journals and conferences. Dr. Santamaria-Navarro has also experience in the private sector for more than 6 years. Prior to his current position, he was a postdoctoral researcher at NASA-JPL and at "Institut de Robòtica i Informàtica Industrial", CSIC-UPC, Barcelona. Dr. Santamaria-Navarro expertise thrives at the intersection between control and perception of mobile robots and his current research interests include state estimation, computer vision and controls, with special focus on unmanned aerial vehicles (UAVs).



**Vincenzo Capuano** received a PhD degree in Microsystems and Microelectronics at École Polytechnique Fédérale de Lausanne (EPFL), a Master of Engineering in Astronautic Engineering from the University of Rome "La Sapienza" and a Bachelor of Engineering (summa cum laude) in Aerospace Engineering from the University of Naples "Federico II". Since 2018, when he joined California Institute of Technology (Caltech) as postdoctoral researcher, his work has focused on the design and development

of innovative and advanced technologies, Vision-based and GPS-based, for Autonomous Spacecraft Navigation. Before, as research scientist (2016-2017) and doctoral assistant (2013-2016) at EPFL he investigated the use of GNSS signals for high earth orbits up to the Moon, contributing to the development and validation of a GPS receiver proof of concept for lunar mission. Dr. Capuano was awarded the Swiss National Science Foundation Advanced and Early Postdoc Mobility Grants, respectively in 2018 and 2019.



**Alexei Harvard** received his B.S. in Mathematics, from the California Institute of Technology (Caltech). He is a researcher at Engineering and Applied Science Department of Aerospace (GAL-CIT). He has several publications in the aerospace field, several related U.S. patents and extensive experience in the industrial sector. His research interests include computer vision and state estimation of spacecraft for space exploration.

of spacecraft for space exploration.



**Amir Rahmani** received his Ph.D in Aeronautics and Astronautics from the University of Washington. He is the supervisor of the Multi-Agent Autonomy group at the NASA Jet Propulsion Laboratory. Amir was an assistant professor of Mechanical and Aerospace Engineering at the University of Miami prior to joining JPL. He has over a decade research experience in distributed space

systems, formation flying, as well as swarm robotics. He is the NASA STTR subtopic manager on coordination and control of swarm of space vehicles.



**Soon-Jo Chung** (M'06–SM'12) received the B.S. degree (summa cum laude) in aerospace engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 1998, and the S.M. degree in aeronautics and astronautics and the Sc.D. degree in estimation and control from Massachusetts Institute of Technology, Cambridge, MA, USA, in 2002 and 2007,

respectively. He is currently Bren Professor of Aerospace and a Jet Propulsion Laboratory Research Scientist in the California Institute of Technology, Pasadena, CA, USA. He was with the faculty of the University of Illinois at Urbana-Champaign (UIUC) during 2009-2016. His research interests include spacecraft and aerial swarms and autonomous aerospace systems, and in particular, on the theory and application of complex nonlinear dynamics, control, estimation, guidance, and navigation of autonomous space and air vehicles. Dr. Chung was the recipient of the UIUC Engineering Deans Award for Excellence in Research, the Beckman Faculty Fellowship of the UIUC Center for Advanced Study, the U.S. Air Force Office of Scientific Research Young Investigator Award, the National Science Foundation Faculty Early Career Development Award, and three Best Conference Paper Awards from the IEEE and the American Institute of Aeronautics and Astronautics. He is an Associate Editor of IEEE TRANSACTIONS ON ROBOTICS, IEEE TRANSACTIONS ON AUTOMATIC CONTROL (since November 2019), and Journal of Guidance, Control, and Dynamics.