Estimation of state, shape, and inertial parameters of space objects from sequences of range images

Matthew D. Lichter*, Steven Dubowsky
Department of Mechanical Engineering, Massachusetts Institute of Technology
Cambridge, MA 02139 USA

ABSTRACT

This paper presents an architecture for the estimation of dynamic state, geometric shape, and inertial parameters of objects in orbit, using on-orbit cooperative 3-D vision sensors. This has application in many current and projected space missions, such as satellite capture and servicing, debris capture and mitigation, and large space structure assembly and maintenance. The method presented here consists of three distinct parts: (1) kinematic data fusion, which condenses sensory data into a coarse estimate of target pose; (2) Kalman filtering, which filters these coarse estimates and extracts the full dynamic state and inertial parameters of the target; and (3) shape estimation, which uses filtered pose information and the raw sensory data to build a probabilistic map of the target's shape. This method does not rely on feature detection, optical flow, or model matching, and therefore is robust to the harsh sensing conditions of space. Instead, it exploits the well-modeled dynamics of objects in space through the Kalman filter. The architecture is computationally fast since only coarse measurements need to be provided to the Kalman filter. This paper will summarize the three steps of the architecture. Simulation results will follow showing the theoretical performance of the architecture.

Keywords: State estimation, shape estimation, parameter estimation, Kalman filter, range images, satellite capture, space debris, large space structures, cooperative sensing

1. INTRODUCTION

Many current and future on-orbit space operations involve physical interaction with dynamic free-floating targets. Examples include the capture and servicing of satellites, the capture and disposal of space debris, and the assembly and maintenance of large space structures[2,14,19]. For these missions, it is critical to have accurate knowledge of the target's motions, shape, and inertial parameters so that interactions with the target can be planned safely. This information is usually unavailable from earth-based sensors and a priori information regarding the shape of the target is often uncertain, especially in the case of damaged satellites, space debris, and warped structure modules. Therefore, current mission concepts expect to use on-orbit 3-D vision sensors to estimate this information[20].

This paper summarizes a method for simultaneously estimating dynamic state, geometric shape, and inertial parameters for an arbitrary dynamic target, using only sequences of range images generated by a team of cooperating sensors. It is assumed that the sensors' relative positions and orientations are accurately known, and that the target is within the field of view of each of the sensors. A priori knowledge about the target is not required, although when available it can be used to speed estimation, reduce sensory requirements, or detect target anomalies (i.e. structural damage). This non-specific need for a priori information makes it well suited to any of the applications described above.

The estimation method consists of three distinct parts, the details of which can be designed independently. The first part consists of kinematic data fusion, a pre-filter which condenses the range image data into a coarse pose estimate at each sample time (see Figure 1). The second part is a Kalman filter, which observes the sequence of pose estimates (surrogate measurements) over time and extracts the full dynamic state and inertial properties of the target using a dynamic model. The final part is a shape estimator, which uses the raw sensory data in conjunction with a refined pose estimate to build a body-centered, probabilistic map of the target's shape.

* lichter@mit.edu; (617) 253-2334; http://robots.mit.edu

Intelligent Robots and Computer Vision XXI: Algorithms, Techniques, and Active Vision,
David P. Casasent, Ernest L. Hall, Juha Röning, Editors, Proceedings of SPIE Vol. 5267
(2003) © 2003 SPIE · 0277-786X/03/$15.00

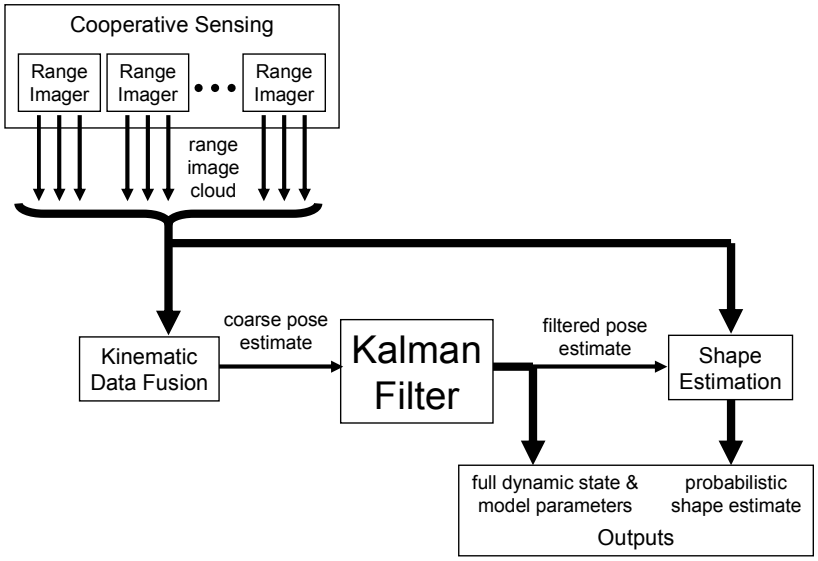


Figure 1: Estimation architecture block diagram.

This approach is powerful in that it takes full advantage of the well-modeled dynamics found in space. It does not rely on feature detection schemes, optical flow, or model matching. For this reason, it is more robust to the harsh sensing conditions of space. Because only rough surrogate measurements are needed for the Kalman filter, the data fusion step can trade accuracy for computational simplicity and speed. Thus, the overall architecture can be very fast. Finally, using accurate sensor models while tracking the body's motions allows the sensors to gain multiple viewpoints of the target and fuse redundant noisy measurements over time in a statistically optimal way.

This paper will describe the details of each part of the estimator, in the context of rigid tumbling targets such as satellites and debris. Emphasis will be given to the Kalman filter design, since its role is central to the architecture and because important design subtleties exist. The paper will also discuss degenerate situations in which estimation becomes difficult (i.e. parameters become unobservable) and potential methods for overcoming them. The paper will conclude with simulation results, showing estimator convergence times, performance versus number of sensors, and performance in degenerate situations.

2. BACKGROUND LITERATURE

The fundamental problem at hand is the simultaneous estimation of both the structure of an object and its motion relative the observers. While many researchers have studied the estimation of shape from known motion (and vice versa), the simultaneous estimation of both is far more difficult. Notable solution approaches are described below.

2.1 Feature-based methods

Numerous methods have been proposed which rely on the continuous tracking of high-level features to determine relative motions of an unknown object or environment. By maintaining an inventory of located features, the methods also estimate high-level geometric structure. Typically a Kalman filter is used to estimate both the feature locations and the motion parameters in a joint framework.

Many feature-based methods have examined the estimation of an unknown object moving with respect to a fixed observer. Natural features of the object (e.g. corners, edges, markers) are located and tracked in time to understand the high-level motions and the general structure of the rigid target. Broida, Chellappa, and Young[3,,22] appear to be the first to employ the Kalman filter for efficient recursive estimation and to employ a mechanics-based physical model of object motions.

Perhaps the most widespread use of feature-based methods in recent years has been in the area of Simultaneous Localization and Mapping (SLAM)[5,13]. Here the task is to construct a feature map of a (usually) static environment while constantly localizing the moving sensors with respect to this map. This is a slight reformulation of the previous problem; the camera is moving instead of the target object.

For many practical applications, however, these methods are not robust because they are dependent upon feature detection. In practice, many phenomena including occlusions, harsh lighting, and reflective materials can make the reliable detection and correspondence of features virtually impossible. Additionally, these methods by themselves do not provide a detailed estimate of the shape of the target; they provide only a sparse set of feature points pertaining to the object. For this reason, they do not fully address the estimation requirements of the applications described in Section 1.

2.2 Pixel-level methods

At the opposite end of the spectrum of methodologies are those which rely on pixel-level information rather than high-level features. Several methods use shape from shading[7], shape from texture[11], optical flow[9], or some combination thereof to compute an estimate of object shape at each time step from monocular cameras. Other methods obtain shape estimates directly from stereo cameras or laser rangers. These shape estimates are effectively merged over time, with relative camera motions estimated between time steps[6,8,17]. The recursive estimation is typically performed via a Kalman filter.

Because of their pixel-level computations, however, these methods are unsuited to the applications here. First, shape computations are highly sensitive to pixel-level detail, which is easily corrupted by the harsh lighting, reflective materials, and highly convoluted surfaces found in space (e.g. wrinkled aluminized Mylar films used for thermal protection on satellites). Second, pixel data is estimated directly in the Kalman state, leading to a very high-dimensional filter implementation. This is computationally expensive and not feasible for the applications involved.

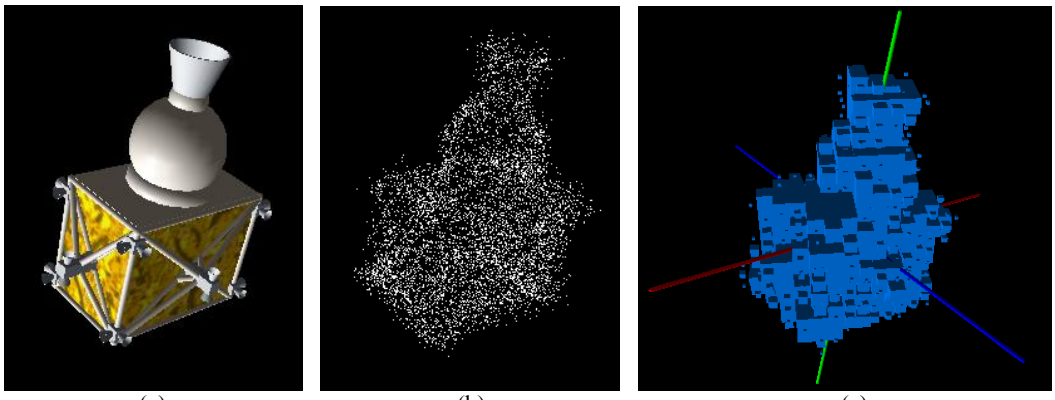
This work attempts to develop an architecture that is not reliant on feature detection schemes or detailed pixel level computations, yet makes full use of well-modeled dynamics and operates in a computationally efficient manner.

3. KINEMATIC DATA FUSION (PRE-FILTERING)

As shown in Figure 1, the first part of the estimation architecture presented here must convert range image data into a rough pose estimate of the target at each sample time. This step is essentially a pre-filter, condensing detailed visual information into high-level kinematic information (surrogate measurements) that can be digested easily by the Kalman filter (Section 4). By the strength of the Kalman filter and the accurate dynamic model, one can afford relatively coarse estimates of pose at this point; therefore this step should focus on computational simplicity and robustness in its processing of the range image data.

Sensory data is assumed to come from a team of range imagers such as stereo cameras or laser-based technologies. The relative kinematics between the sensors is assumed to be known to high accuracy to allow the mosaicking of data amongst sensors. If several sensors are distributed fairly uniformly about the target and have the target within their field of view, then the surface geometry of the target should be captured reasonably well to first order. The sensors' view of the target will be a noisy 3-D cloud of points as shown in Figure 2b.

Finding the centroid of this cloud could provide a rough estimate of target position in space. However, this is not robust since sensors which are very close to the target will provide a higher density and larger number of sample points than sensors located far away. Thus the centroid computations would be biased towards closer sensors. A much better method is to discretize the space into voxels (volume elements) with each voxel having an occupancy level proportional to the number of sample points found within it (see Figure 2c). The occupancy values can then be saturated at some pre-defined threshold to minimize the bias effects of closer sensors.



(a) (b) (c)

Figure 2: (a) Simulated object; (b) simulated range image cloud seen by sensors; (c) voxelized representation of cloud, with principal geometric axes shown.

The location of the geometric centroid of the voxel image is computed to provide a coarse estimate of target position:

$$\vec{r}_m = \frac{\sum n_i \cdot \vec{r}_i}{\sum n_i} \tag{1}$$

where n_i is the occupancy value of the i^{th} voxel, $\vec{r}_i$ is the position of that voxel in sensor coordinates, and $\vec{r}_m$ is the computed centroid of the voxel image. Similarly, second moment information about the voxels is computed, analogous to inertia computations for a solid body:

$$\begin{aligned}
J_{xx} &\equiv \sum n_i \left\{ \left(r_i^{(y)} - r_m^{(y)}\right)^2 + \left(r_i^{(z)} - r_m^{(z)}\right)^2 \right\} & J_{xy} = J_{yx} &\equiv -\sum n_i \left(r_i^{(x)} - r_m^{(x)}\right)\left(r_i^{(y)} - r_m^{(y)}\right) \\
J_{yy} &\equiv \sum n_i \left\{ \left(r_i^{(x)} - r_m^{(x)}\right)^2 + \left(r_i^{(z)} - r_m^{(z)}\right)^2 \right\} & J_{yz} = J_{zy} &\equiv -\sum n_i \left(r_i^{(y)} - r_m^{(y)}\right)\left(r_i^{(z)} - r_m^{(z)}\right) \\
J_{zz} &\equiv \sum n_i \left\{ \left(r_i^{(x)} - r_m^{(x)}\right)^2 + \left(r_i^{(y)} - r_m^{(y)}\right)^2 \right\} & J_{zx} = J_{xz} &\equiv -\sum n_i \left(r_i^{(z)} - r_m^{(z)}\right)\left(r_i^{(x)} - r_m^{(x)}\right)
\end{aligned} \tag{2}$$

where the superscripts on r_i and r_m denote the x, y, or z component of those vectors. The second moments are assembled into a matrix whose eigenvectors are computed, satisfying

$$[J] \equiv \begin{bmatrix} J_{xx} & J_{yx} & J_{zx} \\ J_{xy} & J_{yy} & J_{zy} \\ J_{xz} & J_{yz} & J_{zz} \end{bmatrix} = [R_m] \cdot [\Lambda] \cdot [R_m]^T \tag{3}$$

Here $[J]$ is the geometric moment matrix (analogous to an inertia matrix), $[\Lambda]$ is the diagonal eigenvalue matrix of $[J]$, and $[R_m]$ is the eigenvector matrix. This computation locates the principal geometric axes of the voxel image. $[R_m]$ is thus a rotation matrix describing the attitude of the principal geometric axes with respect to the sensor reference frame (see Figure 2c).

The centroid position $\vec{r}_m$ and rotation matrix $[R_m]$ thus represent the output of the kinematic data fusion step; they are the surrogate measurements of target position and attitude which feed into the Kalman filter. If the target is rigid, the voxel representation is approximately constant in shape, and thus tracking the principal geometric axes provides a simple way to coarsely track the target's pose. It is important to note that these tracked axes do not correspond to the actual center of mass or principal inertial axes of the target; they are based merely on superficial geometry of the body.

The computationally simple method presented here has important degeneracies, namely when the target has a high degree of axial symmetry. In such cases, two or three of the eigenvalues of [*J*] will have approximately the same value, which leads to the loss of unicity in the solution of Equation 3. In other words, a long cylinder has only one obvious geometric axis (the other two can be specified arbitrarily as long as they form an orthogonal basis), and a sphere has no obvious geometric axes. This minor problem can be handled by introducing more information (higher-order moments, color information from the sensors, etc.) to constrain the solution of Equation 3.

It should be noted that there are many possible ways to provide a surrogate measurement to the Kalman filter. The method presented here is computationally fast, easy to implement, and proves to be quite adequate in many situations (see Section 6). It is used here as an illustrative example to demonstrate the utility of the architecture as a whole.

4. KALMAN FILTERING

The Kalman filter forms the core of the estimation architecture, using the surrogate measurements along with an accurate dynamic model to estimate the full dynamic state and inertial parameters of the target. The dynamic state consists of rotational and translational positions and velocities. External forces and torques on the target are assumed to be negligible. Gravity gradient torques and orbital mechanics can be incorporated into the model; however their contribution is negligible over short time intervals and in practice it is sufficient to model them as process noise in the Kalman filter.

The parameters to be estimated include the principal inertias of the target (relative magnitudes only), and the kinematics of the principal inertial axes and center of mass with respect to the principal geometric axes (surrogate measurement). As discussed in Section 3, the reference frame observed by the data fusion portion does not correspond to the principal inertial axes of the target (see Figure 3). However, since both sets of axes are fixed to the body, the relative kinematics between the frames is constant and can be parameterized and estimated by the Kalman filter.

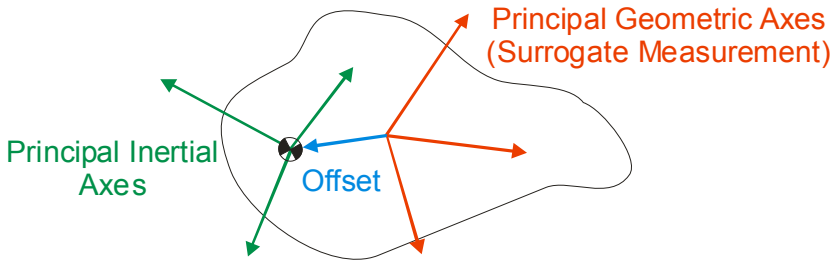


Figure 3: Body-fixed reference frames on the target.

4.1 Rotational estimation

Because the rotational and translational dynamics of the target are decoupled, the estimation can be performed using two separate Kalman filters. For rotational estimation, the target object is assumed to be rigid and external torques on the body are assumed to be negligible, which provides a good model for most situations. Rotations are represented not using rotation matrices, but with unit quaternions (Euler parameters). This simplifies many of the equations and provides a near-minimal framework for dealing with rotations. Additionally, the unit quaternion has only a single normality constraint, which is much easier to enforce than the orthonormality of rotation matrices. Euler angles are not considered due to their well-known singularity problems. For this problem, the quaternion representation is sufficient and robust.

To be estimated are the rotational velocities in body coordinates ($\vec{\omega}_b$), the attitude quaternion of the principal inertial axes ($\vec{q}_b$), the principal inertias of the target ($\vec{I}$), and the rotational offset between the surrogate measurement and the principal inertial axes ($\vec{q}_d$). Only relative magnitudes of the inertias are observable, so $\vec{I}$ is normalized. These quantities follow the dynamic model

$$\frac{d}{dt}\begin{Bmatrix} \omega_{b1} \\ \omega_{b2} \\ \omega_{b3} \\ q_{b0} \\ q_{b1} \\ q_{b2} \\ q_{b3} \\ \vec{I} \\ \vec{q}_d \end{Bmatrix} = \begin{Bmatrix} \frac{I_2 - I_3}{I_1}\omega_{b2}\omega_{b3} \\ \frac{I_3 - I_1}{I_2}\omega_{b3}\omega_{b1} \\ \frac{I_1 - I_2}{I_3}\omega_{b1}\omega_{b2} \\ \frac{1}{2}\left(-\omega_{b1}q_{b1} - \omega_{b2}q_{b2} - \omega_{b3}q_{b3}\right) \\ \frac{1}{2}\left(\omega_{b1}q_{b0} - \omega_{b2}q_{b3} + \omega_{b3}q_{b2}\right) \\ \frac{1}{2}\left(\omega_{b1}q_{b3} + \omega_{b2}q_{b0} - \omega_{b3}q_{b1}\right) \\ \frac{1}{2}\left(-\omega_{b1}q_{b2} + \omega_{b2}q_{b1} + \omega_{b3}q_{b0}\right) \\ \vec{0} \\ \vec{0} \end{Bmatrix} \tag{4}$$

This represents the dynamics of a rigid body under torque-free motion[15] and specifies the parameters to be estimated as constants. The relationship between the surrogate measurement and the principal axes is given by

$$\vec{q}_m = \vec{q}_b \otimes \vec{q}_d \tag{5}$$

where $\vec{q}_m$ is the quaternion parameterization of the surrogate measurement [R_m] and the operator $\otimes$ is used to signify quaternion multiplication[15]. Equation 5 merely states that a constant rotation exists between the surrogate measurement and the principal inertial axes of the body.

Equations 4 and 5 represent the dynamic and measurement models for the rotational Kalman filter. These equations present several challenges: (1) they are highly nonlinear; (2) the quaternions and inertia vectors contain normality constraints; and (3) the sum of any two inertia components must be greater than the third (by natural physical bounds). Further, the state estimation errors are kinematically correlated; for example, errors in the estimation of the offset quaternion are not independent of errors in the estimation of the principal inertia axes' attitude.

Many methods exist to handle these conditions[1]; however a full discussion is beyond the scope of this paper. The methods used here include an unscented Kalman filter to handle the nonlinear models[10,21], quaternion multiplications rather than vector additions during the Kalman update to maintain the quaternion normalities[12], re-parameterization of the quaternions to achieve minimum-order covariance structures, re-parameterization of the inertia vector to an unbounded space, and coordinate changes to kinematically de-correlate estimation errors in the state vector.

4.2 Translational estimation

Translational estimation is much simpler than rotational estimation. To be estimated is the translational velocity ($\vec{v}_b$) and position ($\vec{r}_b$) of the center of mass, and the translational offset between the surrogate measurement and the center of mass ($\vec{r}_d$) (in a body-fixed coordinate frame). In the absence of external forces, the dynamic model is given by

$$\frac{d}{dt}\vec{v}_b = \vec{0} \qquad \frac{d}{dt}\vec{r}_b = \vec{v}_b \qquad \frac{d}{dt}\vec{r}_d = \vec{0} \tag{6}$$

The measurement model is given as

$$\vec{r}_m = \vec{r}_b + [R_m]\cdot\vec{r}_d \tag{7}$$

where [R_m] is the rotation matrix from the surrogate measurement. It is usually more appropriate to use a filtered version of [R_m] from the rotational filter rather than using the much noisier surrogate measurement directly. While this does introduce slight coupling between the filters, the performance tends to be superior in practice.

Equations 6 and 7 are linear and thus a basic (linear) Kalman filter[4] provides excellent results. These results will be discussed in Section 6.

4.3 Degenerate conditions

There are a few degenerate conditions of the target which should be noted. The first occurs when two or more principal inertias are identical. In this scenario, the principal axes are undefined by definition, since an infinitude of orientations of the principal frame yield a diagonal inertia matrix. In other words, there is no unique solution to the estimation problem. However, this does is not a problem since the Kalman filter will converge on one possible solution and remain there.

If the target has no rotational velocity, then the inertias and the kinematic offset are unobservable. As a result, the attitude of the principal inertial axes will be unknown. However, a filtered version of the surrogate measurement will be available and therefore the position of that target-fixed reference frame can still be tracked.

If the target is spinning about a single axis (a flat spin), inertias are again unobservable. However, not all inertia information is lost. The major or minor principal inertial axis must be coincident with the angular velocity vector, and the center of mass must lie on this axis. This vector is observable in global coordinates. Again, the attitude of the principal inertial axes will be uncertain, but the filtered version of the surrogate measurement will be available to track the object.

From a practical standpoint, the lack of inertia information in the second and third cases is not a grave concern. Inertia information will be available if and only if it is needed to predict the target trajectory, which is the ultimate goal to the mission planner. Also, capturing a docile target might be a preferred scenario to capturing a wildly tumbling object, so inducing a multi-axis tumble in the target by external means in order to gain inertia information might be unwise in the long run.

5. SHAPE ESTIMATION

With accurate knowledge about the trajectory of the target, it becomes possible to fuse incoming raw sensory data into a body-fixed map of the target. This is a classic stochastic mapping problem, where the environment (target) motions are known with respect to the sensors. Again, numerous methods can be used to do this[16,18]. This paper will illustrate one simplistic approach.

At each sample time, a probability density function (PDF) in 3-D space can be generated to stochastically describe the estimated shape of the target. This overall PDF can be obtained by convolving each sensor data point with the PDF describing the noise characteristics of the sensor. That is

$$p_{shape}(\vec{r}) = \sum_i \delta(\vec{r} - \vec{r}_i) * p_{sensor}(\vec{r}) \tag{8}$$

where $p_{shape}(\vec{r})$ is the PDF describing the relative likelihood that the target surface exists at point $\vec{r}$, $p_{sensor}(\vec{r})$ is the PDF describing the noise distribution on each sensor data point, $\vec{r}_i$ is the location of the i^{th} data point, $\delta(\ )$ is the Dirac delta function, and * is the convolution operator. The goal of shape estimation is to accurately depict $p_{shape}(\vec{r})$.

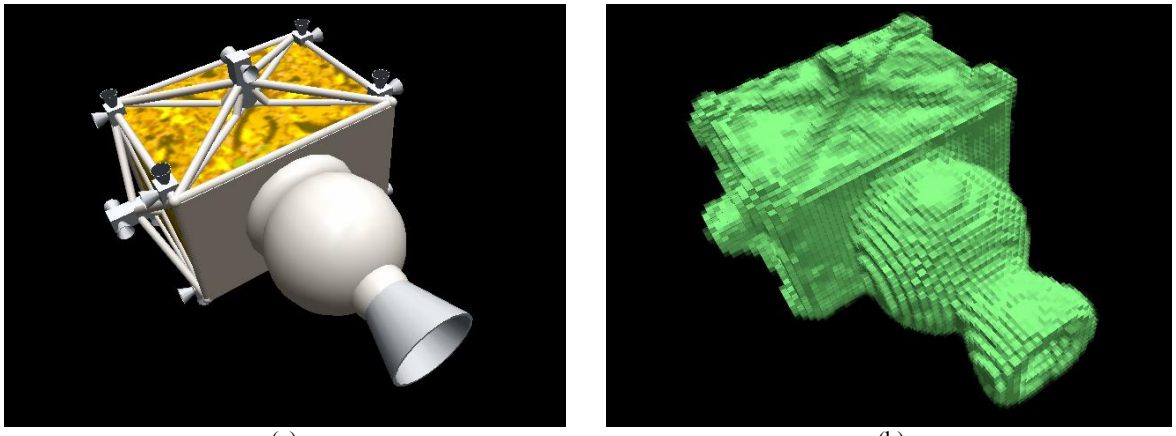
It is important that the coordinate system used in Equation 8 be a body-fixed reference frame, so that data can be fused across time as the target moves. It is best to use a filtered version of the surrogate measurement frame, rather than the principal inertia frame (see Figure 3), because it can be tracked even under degenerate conditions (see Section 4.3).

Equation 8 can be computed at each sample time to obtain a PDF based on the current data. This should be combined somehow with the PDFs from all previous sample times to yield a more accurate overall PDF. A computationally efficient method would be recursive, making incremental improvements to its estimate using new information. One solution would be to compute the new estimate as a weighted average of the previous estimate and the new information. This amounts to applying a forgetting factor to old data:

$$p_{shape}^{1\to t} = \alpha \cdot p_{shape}^{1\to(t-1)} + (1-\alpha) \cdot p_{shape}^{t} \tag{9}$$

where $p_{shape}^{1\to t}$ is the cumulative shape estimate incorporating information from time samples 1 through t, $p_{shape}^{1\to(t-1)}$ is the old cumulative estimate, p_{shape}^{t} is the PDF based strictly on the current sensory data, and α is a forgetting factor between 0 and 1. The larger the forgetting factor, the slower old information is forgotten. The cumulative estimate can be initialized to a priori knowledge of the target shape, if it exists.

In practice, it is generally much easier to deal with discrete sums than integrals in the convolution process. Therefore, the body-fixed space is discretized into voxels, this time with a much finer resolution than that used in Section 3. The finer resolution is enabled because data is being used from all the sample times rather than a single sample, therefore much more information is available. Figure 4 shows the resultant voxel estimate of a simulated target based on 100 sample times, with voxels having a probability of occupancy higher than 50 percent shown.



(a) (b)

Figure 4: (a) Simulated target; (b) voxel-based shape estimate.

This is a rather rudimentary approach but demonstrates the concept of the architecture. Clearly, more sophisticated approaches could be applied to this problem if necessary. This is an area of ongoing research.

6. RESULTS

Computer simulations were used for the preliminary study of the estimation architecture. Representative space objects were built into a virtual environment. Simulated sensors were placed within this environment and their range image data was synthesized (see Figure 5). Sensors were given a field of view of 45 degrees and a resolution of 200 by 200 pixels. Generally, about one to two thousand data points from the target were provided by each sensor. Gaussian white noise was added to the simulated data, with standard deviation in the range measurement being about 1 percent of the measurement itself, and standard deviation in the focal plane measurements being one-half pixel.

(a) (b)

Figure 5: (a) Simulated environment; (b) synthesized range image.

6.1 Kinematic data fusion

Studies were performed to evaluate surrogate measurement coarseness as a function of the number of cooperative sensors used. Clearly, using more sensors provides better coverage of data over the target surface by providing additional vantage points. In a Monte Carlo simulation involving several thousand trials, a simulated target was placed in random positions and orientations, whereupon range data was synthesized and kinematically fused to produce surrogate measurements of the target pose. The sensors were distributed uniformly in space about the target, and their number was varied across trials. The coarse surrogate measurements were compared to the actual pose of the target and the errors were quantified. Figure 6 shows the mean attitude and position errors pertaining to the simulated target shown in Figure 4a. Several other simulated targets were used with similar results.

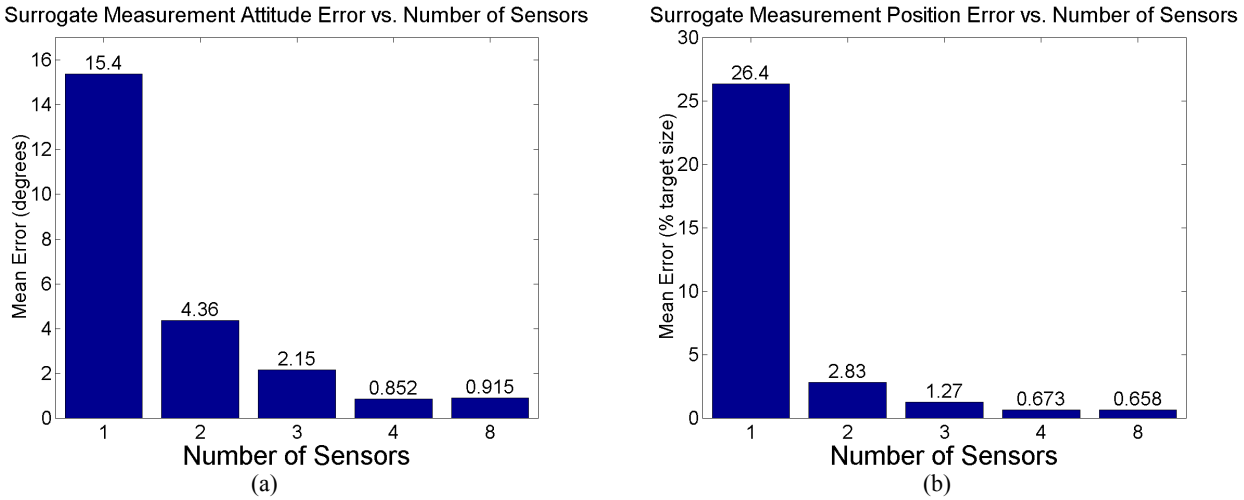


Figure 6: Surrogate measurement errors vs. number of sensors: (a) attitude error; (b) position error.

These results suggest the reasonable accuracy obtainable from the computationally simple algorithm presented in Section 3. With multiple sensors, surrogate measurements are accurate to the order of a few degrees and a few percent of the target size, which is sufficient for the Kalman filter. These results also show the diminishing returns of using more than about four sensors. The noise in the sensors and the artifacts of the discrete sampling and voxelization processes mean that some error in the surrogate measurement will always exist. Finally, these results are a reminder that the method described here is inappropriate for use with a single sensor, as data is heavily biased toward the sensor and absent from the back side of the target. A slightly different data fusion method would be needed if only one sensor were available.

6.2 Kalman filtering

Computer simulations were again used to study the Kalman filter portion of the estimator. For these studies, four sensors were used, distributed uniformly about the target in a tetrahedral formation. The data fusion method as described above was used to provide a surrogate measurement to the filter.

Figure 7 shows typical convergence plots for the estimator for a non-degenerate target. The sampling rate of the sensors need only be fast enough to prevent aliasing of the surrogate measurement as the target tumbles. One can see that the dynamic state and parameters are estimated rather quickly; usually less than ten target rotations need to be observed.

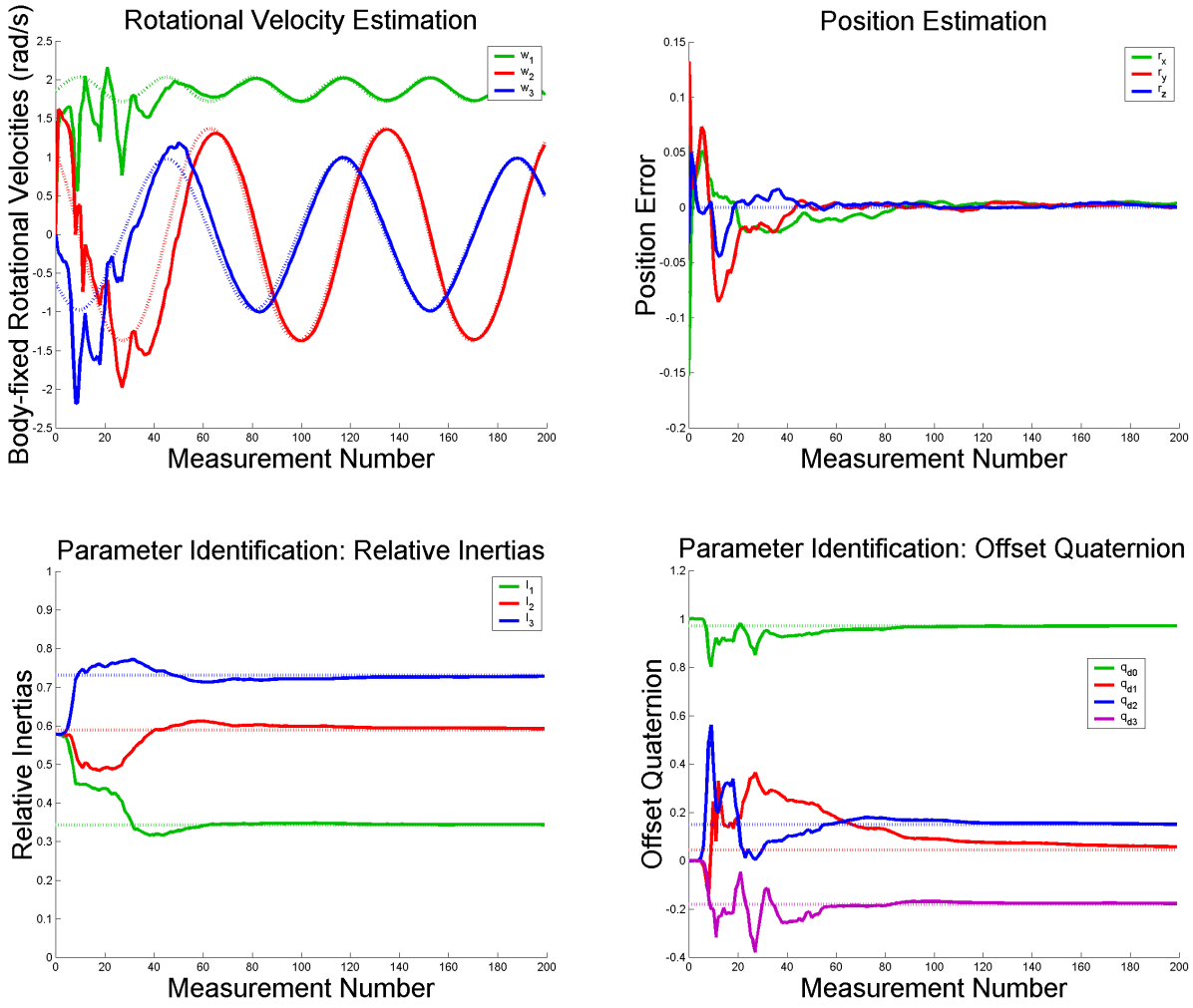


Figure 7: Kalman filter convergence plots.

As the tumble conditions of the target approach the degeneracies (single-axis spins), convergence in the Kalman filter becomes more difficult. In a Monte Carlo simulation involving 2500 trials, the initial spin condition of the target was varied, and the time to convergence of the filter was recorded. Convergence was defined as estimating the attitude of the principal inertial axes to within 10 degrees and estimating the principal inertia vector to within 3 degrees of the true value (these were generally the slowest values to converge). The principal inertias of the target were defined to be $[0.3431 \quad 0.5893 \quad 0.7314]^T$, a value well away from the inertia equality degeneracy. Initial angular momentum was held constant among all trials so that rotation speed would not factor into the convergence times. Figure 8 is a scatter plot showing Kalman filter performance versus the spectrum of spin conditions of the target. The spin condition is parameterized by the nutation angle, the angle between the major principal inertial axis and the momentum vector at the instant of zero spin about the intermediate principal axis. Thus, a nutation angle of 0 degrees indicates a flat spin about the major principal axis, 90 degrees indicates a flat spin about the minor principal axis, and an intermediate angle indicates a multi-axis tumble.

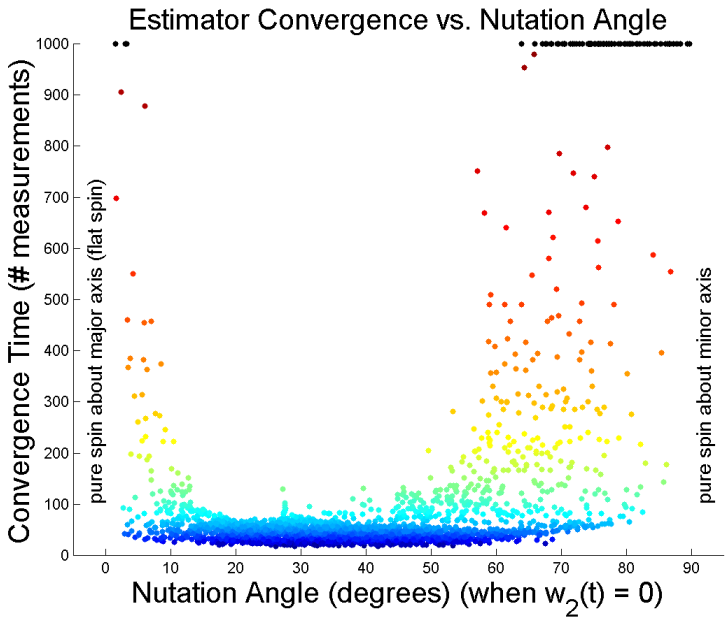


Figure 8: Kalman filter convergence across the spectrum of target spin conditions.

Figure 8 affirms that estimation is slow near degenerate spin conditions and fast for targets in a multi-axis tumble. The filter could not converge in the allotted time for many targets in near-flat spins, as shown by the dark band of points along the top of the plot. Targets in a multi-axis tumble have rich information content embedded in their trajectory while flatly spinning targets have little. As discussed in Section 4.3, this may not be a great concern, since targets in flat spins still have predictable trajectories and are generally regarded as docile compared to targets in a multi-axis tumble.

6.3 Computation speed

No explicit studies were performed to evaluate architecture speed. However, virtually all the computational burden in the simulations came from the graphics and the synthesis of sensory data, and the estimator was still able to operate at a rate of several Hz. This suggests that the architecture would be limited by the data rate of the sensors in practice, not by internal computations. Also recall that sensor sampling rate only needs to be fast enough to prevent aliasing of the surrogate measurement as the target tumbles.

7. CONCLUSIONS

This paper has presented a new approach for the estimation of dynamic state, geometric shape, and inertial parameters of a space target using range data from a team of sensors. At its core is a Kalman filter, which exploits the high-fidelity dynamic model of objects in the space environment. Simple kinematic data fusion serves as a pre-filter for the raw sensory data, condensing it into high-level information that the Kalman filter can use easily. With a filtered estimate of body pose, shape estimation reduces to a stochastic mapping problem where relative motions between sensors and target are known. This method is robust in that it does not rely on feature detection, model matching, or pixel-level computations and thus may be well suited to the harsh visual sensing conditions of space.

Simulation results suggested the utility and performance of this architecture. Surrogate measurements from the kinematic data fusion step were shown to be sufficiently accurate for input to the Kalman filter. The Kalman filter was shown to converge rapidly under non-degenerate conditions. Targets in flat spins were shown to have low parameter observability compared to targets in multi-axis tumbles, however, such degenerate targets still had predictable trajectories. Finally the architecture was found to be fast and would probably be limited by sensor speed rather than computation power in practice. Research on this project is continuing, using experimental hardware to study the architecture using real sensory data.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the Japanese Aerospace Exploration Agency (JAXA) for their support of this work. The authors would also like to thank Dr. Vivek Sujan, Dr. Noam Brook, and Dr. Karl Iagnemma for their helpful dialogue during the development of this work.

REFERENCES

1. Bard, J., *Nonlinear Parameter Estimation*, New York: Academic Press, 1974.
2. Bornschlegl, et al, "Space Robotics in Europe, a Compendium," *Proceedings of the 7th International Symposium on Artificial Intelligence, Robotics, and Automation in Space: i-SAIRAS 2003*, NARA, Japan, May 19-23, 2003.
3. Broida, T.J. and R. Chellappa, "Estimating the Kinematics and Structure of a Rigid object from a Sequence of Monocular Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 6, pp. 497-513, 1991.
4. Brown, R.G., and P.Y.C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, 2nd Ed., New York: John Wiley & Sons, 1992.
5. Dissanayake, M.W.M.G., et al, "A Solution to the Simultaneous Localization and Map Building (SLAM) Problem," *IEEE Transactions on Robotics and Automation*, Vol. 17, Issue 3, pp. 229-241, Jun 2001.
6. Heel, J., "Direct Dynamic Motion Vision," *Proceedings of the IEEE International Conference on Robotics and Automation*, Vol. 2, pp. 1142-47, May 13-18, 1990.
7. Horn, B.K.P. and M.J. Brooks, eds., *Shape from Shading*, Cambridge, Massachusetts: MIT Press, 1989.
8. Horn, B.K.P. and J.G. Harris, "Rigid Body Motion from Range Image Sequences," *Computer Vision, Graphics, and Image Processing (CVGIP): Image Understanding*, Vol. 53, No. 1, Jan, pp. 1-13, 1991.
9. Horn, B.K.P. and B.G. Schunck, "Determining Optical Flow," *AI* , Vol. 17, No. 1-3, pp. 185-203, August 1981.
10. Julier, S.J., and J.K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," *Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, Orlando, Florida, 1997.
11. Kanatani, K. and T. Chou, "Shape from Texture: General Principle," *AI*, Vol. 38, pp. 1-48, 1989.
12. Lefferts, E.J., F.L. Markley, and M.D. Shuster, "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, Sept.-Oct. 1982, pp 417-429.
13. Leonard, J.J. and H.F. Durrant-Whyte, *Directed Sonar Sensing for Mobile Robot Navigation*, Boston, MA: Kluwer Academic, 1992.
14. Piedboeuf, J-C. and É. Dupuis, "Recent Canadian Activities in Space Automation & Robotics – An Overview," *Proceedings of the 7th International Symposium on Artificial Intelligence, Robotics, and Automation in Space: i-SAIRAS 2003*, NARA, Japan, May 19-23, 2003.
15. Sidi, M.J., *Spacecraft Dynamics and Control: a Practical Engineering Approach*, New York: Cambridge University Press, 1997.
16. Smith, R.C., and P. Cheeseman, "On the Representation and Estimation of Spatial Uncertainty," *International Journal of Robotics Research*, vol. 5, no. 4, pp. 56-68, 1987.
17. Soatto, S., R. Frezza, and P. Perona, "Motion Estimation via Dynamic Vision," *IEEE Transactions on Automatic Control*, Vol. 41, No. 3, March 1996, pp. 393-413.
18. Sujan, V., and S. Dubowsky, "Visually Built Task Models for Robot Teams in Unstructured Environments," *Proceedings of 2002 IEEE International Conference on Robotics and Automation (ICRA 2002)*, Washington, D.C., May 11-May 15, 2002.
19. Ueno, H., et al, "Space Robotic Mission Concepts For Capturing Stray Objects," *Proceedings of the 23rd International Symposium on Space Technology and Science*, Matsue, Japan, May 26th-June 2nd, 2002.
20. Wakabayashi, Y., et al, "A Compact Laser Range Finder for Space Applications," *Proc. SPIE Vol. 3714: Enabling Photonic Technologies for Aerospace Applications*, pp. 131-138, A.R. Pirich, E.W. Taylor, eds., 1999.
21. Wan, E.A., and R. van der Merwe, "The Unscented Kalman Filter," *Kalman Filtering and Neural Networks*, Edited by Simon Haykin, New York: Wiley Publishing, 2001, pp. 221-280.
22. Young, G-S.J. and R. Chellappa, "3-D Motion Estimation Using a Sequence of Noisy Stereo Images: Models, Estimation, and Uniqueness Results," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 8, pp. 735-759, Aug 1990.