JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 34, No. 1, January–February 2011

Engineering Notes

Heuristic Pose Estimation of a Passive Target Using a Global Model

D. S. Wokes* and P. L. Palmer†
Surrey Space Centre,
Guildford, England GU2 7XH, United Kingdom

DOI: 10.2514/1.51030

I. Introduction

MANY eyes are focused upon the objective of quickly obtaining the position and orientation of a target object from a single image or a stream of images. This topic has endured over past and more recent decades (see Besl and Jain [1] or Lowe [2], who study object reconstruction, and Ruel et al. [3], who develop real-time pose estimation techniques) but remains problematic to implement in more generic scenarios. In both air and space environments, scenarios arise that require swift automated guidance with resource-constrained systems: for example, missile guidance (see Davidovitz and Shinar [4]). Sometimes, the necessity for speed outweighs the desire for precision (see Hmam and Kim [5], Agarwal and Chaudhuri [6], and Beaulieu [7]).

Consistent throughout pose determination methods is the requirement of a sufficiently accurate estimate to begin the process with. The proposed algorithms for International Space Station robotic arm operations from Stieber et al. [8] are dependent on an initial pose estimate. Edge-line correspondence methods are recognized for lacking speed and adaptivity (see Shakunaga [9]). Lowe [2,10] also attempted edge-line correspondence methods to determine the pose of a target, both for static and time-evolving images. This work evolved further with Cropp [11], by solving the edge-line correspondence problem with a RANSAC algorithm (see Fischler and Bolles [12]). The approach suffers from the same constraint: without an initial pose estimate from another system, the pose estimate could take almost 180 times longer than the desired image refresh rate.

Heuristics are challenging for an autonomous device but can unlock the ability and potential for more complex and more accurate pose determination techniques that would otherwise take an unacceptably long time. This Note introduces a novel way of filling in the heuristic gap between image acquisition and edge-line correspondence/specific target recognition. Under the assumption that the target's dimensions are known to an acceptable level of accuracy, it is possible to create a spheroid to model that target. A spheroid is defined as a degenerate ellipsoid: there exists an axis of symmetry for the spheroid. This spheroid describes, *up to its axis of symmetry*, the target's position and attitude. From a single image the bounding ellipse of the target's projection can be extracted, from which the descriptive spheroid can be reconstructed and a pose estimate for the target object can be deduced. Applying a heuristic pose estimation method attempts to create a computational shortcut to more accurate techniques, sacrificing speed for accuracy.

Received 3 June 2010; revision received 27 August 2010; accepted for publication 12 October 2010.

*Researcher; D.Wokes@googlemail.com. Member AIAA.
†Reader; P.Palmer@surrey.ac.uk. Member AIAA.

In the next section the spheroid reconstruction process is reviewed. An example method of obtaining the spheroid model to describe a target object is described in Sec. III, which minimizes the errors obtained in the approximation. In Sec. IV, this method is used to determine a descriptive spheroid for a proposed satellite of the University of Surrey. In Sec. V, the heuristic theory is applied to this model for varying radial distances. The value of describing a target globally is encapsulated by comparing the reconstruction accuracy with a feature-detection technique, the findings of which are summarized and discussed in Sec. VI.

II. Spheroid Reconstruction Overview

In this Note it is assumed that the target's dimensions are known, which enables a spheroidal description of the target to be made before rendezvous/interception. If the target is described as a spheroid, then its perspective projection is given by an ellipse (see Eberly [13]), from which the spheroid can then be reconstructed. Wokes and Palmer derive the reconstruction of a spheroid, of known size, from its perspective projection onto an image plane [14] and image sphere [15]. In this section, the governing equations for the reconstruction process are reviewed.

The advantage of considering an image sphere rather than a plane comes from a simpler, analytic reconstruction process and methods of correcting image sphere ellipses when their solutions are invalid. The conversion from an image plane ellipse to an image sphere ellipse is given in the above citations.

A. Defining the Projection Space

A spheroid is a degenerate ellipsoid; it has an axis of symmetry. Such a geometric surface is described with a position vector $\underline{q}$ from the camera's focal point to the spheroid's center, an orientation vector $\underline{p}$ of its axis of symmetry, a half-length A along its axis of symmetry, and a half-length B along its degenerate axes. The vectors $\underline{q}$ and $\underline{p}$ are defined in the camera's coordinate frame S_c: a Cartesian frame centered at the focal point of the projection. With an image sphere being considered, these Cartesian axes are arbitrarily chosen: define the x axis to lie parallel to the focal axis, the z axis to point toward the north pole, in terms of the image sphere coordinates (defined below) and the y axis to complete the orthogonal set.

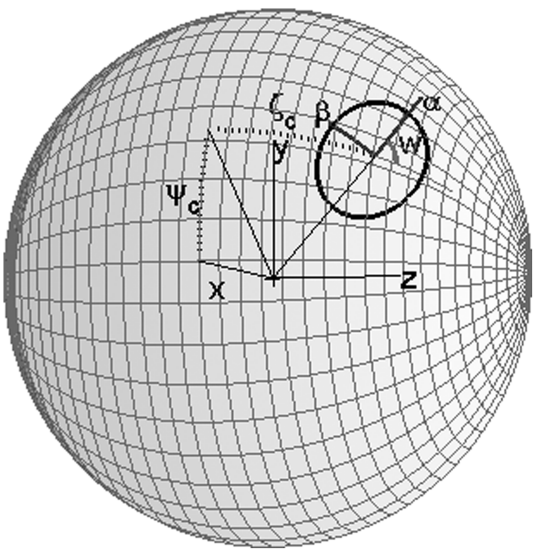


Fig. 1 Image sphere ellipse.

The description of an ellipse on a unit sphere is defined entirely with angles; the semimajor axis is described not with a length, but instead by an arc length α and the semiminor axis described by an arc length β. Based in the camera coordinate system, define ζ_c as the angle between the focal axis and the line of latitude on which the ellipse center is, ψ_c as the angle between the focal axis and the line of longitude on which the ellipse center is. Finally, define w as the angle from the semimajor axis of the ellipse to that line of longitude. In essence, ζ_c is describing the latitude, and ψ_c is describing the longitude (see Fig. 1). Let S_e denote the image sphere ellipse coordinate frame, centered at the focal point, with the first axis aligned pointing through the ellipse center and the second and third axes aligned with the ellipse's semiminor and semimajor axes, respectively. There is a rotation matrix V: $S_e \to S_c$, given by

$$V \doteq \begin{pmatrix} \cos\psi_c \cos\zeta_c & -\sin\psi_c \cos w + \cos\psi_c \sin\zeta_c \sin w & -\sin\psi_c \sin w - \cos\psi_c \sin\zeta_c \cos w \\ \sin\psi_c \cos\zeta_c & \cos\psi_c \cos w + \sin\psi_c \sin\zeta_c \sin w & \cos\psi_c \sin w - \sin\psi_c \sin\zeta_c \cos w \\ \sin\zeta_c & -\cos\zeta_c \sin w & \cos\zeta_c \cos w \end{pmatrix} \tag{1}$$

The image sphere ellipse described in the camera coordinate frame is shown in Fig. 1. The described ellipse is that which lies between the spheroid and the focal point of the projection, *not* the ellipse that appears behind the focal point.

B. Reconstructing a Spheroid from an Image Sphere Ellipse

Assuming a prolate spheroid is being viewed, then the equations for the reconstructed pose are given as follows: define γ_i and δ_i as

$$\gamma_1 \doteq \cos\alpha\sqrt{A^2 + B^2\cot^2\beta} \qquad \gamma_2 \doteq \sin\alpha\sqrt{A^2\cot^2\alpha - B^2\cot^2\beta}$$

$$\delta_1 \doteq \sec\beta \qquad \delta_2 \doteq \sqrt{\tan^2\alpha - \tan^2\beta}$$

Then

$$\underline{q} = \gamma_1\delta_1\underline{V}_{(1)} + \vartheta_1\gamma_2\delta_2\underline{V}_{(3)} \tag{2}$$

$$\underline{p} = \vartheta_2\left\{\frac{\gamma_2\delta_1}{\sqrt{A^2 - B^2}}\underline{V}_{(1)} + \vartheta_1\frac{\gamma_1\delta_2}{\sqrt{A^2 - B^2}}\underline{V}_{(3)}\right\} \tag{3}$$

where $\vartheta_i = \pm 1$; there are two solutions to both the position and orientation vectors, which are reflections of each other through the $\underline{V}_{(1)}$–$\underline{V}_{(2)}$ plane. This is shown in Fig. 2. If the spheroid is oblate rather than prolate, then switching $\underline{V}_{(3)}$ with $\underline{V}_{(2)}$, α with β, and A with B obtains the reconstruction equations for the $B > A$ case.

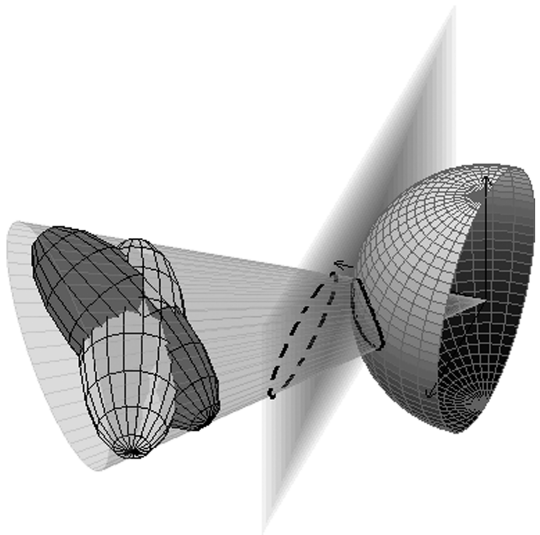


Fig. 2 Spheroid projected onto a plane (dashed ellipse) and a sphere (solid ellipse). The reconstruction yields both the original spheroid and its dual solution.

The example satellite considered in this Note is modeled by a prolate spheroid, where Eqs. (2) and (3) are used to reconstruct their pose estimates. It should be noted that ambiguity remains in the sign of ϑ_2, as well as the sign of ϑ_1. This is a consequence of modeling the target as a conic with a plane, as well as an axis, of symmetry. The axial symmetry means that the roll of the target (assuming this is the axis of symmetry for the descriptive spheroid) is also undetermined.

C. Ellipse Correction Algorithms

Considering Eqs. (2) and (3), under certain circumstances spheroid reconstruction may fail. This is due to the value of $\sqrt{A^2\cot^2\alpha - B^2\cot^2\beta}$ becoming imaginary and yielding complex solutions to the position and orientation. Thus, the constraint placed upon the image sphere ellipse parameters for a real reconstruction, can be rewritten as

$$\frac{A}{B} \geq \frac{\tan\alpha}{\tan\beta} \tag{4}$$

In essence, this equation puts an upper bound on the eccentricity of the image sphere ellipse. By considering a sphere rather than a spheroid, the intuitive constraint becomes $\alpha = \beta$. Choosing how to correct the ellipse, however, depends on the scenario considered. If an image is blurred (for example, missile tracking), the reduction of α may be appropriate. If there is confidence in determining the pixels associated with the target, increasing β may be appropriate. If the desire is to find the solution with minimal change to α and β, a combination of both.

In the reconstruction scenarios considered, only faint background noise is added, pertaining to a high confidence in determining the pixels representing the projection of the target. The target is modeled as both a spheroid and a sphere in order to compare accuracies; when modeled as a sphere, α is reduced to match β. This choice is discussed in Sec. IV. When modeled as a spheroid, the following equations are used:

$$\alpha_{\text{corrected}} \doteq \alpha_{\text{detected}} \tag{5}$$

$$\beta_{\text{corrected}} \doteq \tan^{-1}\left(\frac{B\tan\alpha_{\text{detected}}}{A}\right) \tag{6}$$

where α_{detected} is the original arc length and $\alpha_{\text{corrected}}$ is the modified arc length (likewise for β). Applying these expressions leads to the result of $\underline{p} \perp \underline{q}$: a pose that would yield the highest possible eccentricity with its projection. Note that without changing to a projection onto a sphere, this analysis would not have been possible.

III. Choosing the Descriptive Spheroid

When a nonspheroidal object is modeled as a spheroid, errors are obtained in its reconstruction from an ellipse. Even for a cylinder it can be shown that if the ellipse algorithm is minimum bounding, there does not exist a spheroid that will always yield the correct position when reconstructed from its projection. The objective of this section is to describe a method for choosing the descriptive spheroid (its dimensions, center, and orientation in the target's body frame) that minimizes these errors.

Unlike the spheroid reconstruction process, it is assumed that choosing the descriptive spheroid does not have the same time constraints. To create this cost function, a series of projections of the target model are generated and their bounding ellipses found. The seven variables describing the descriptive spheroid are then chosen to minimize the error in the target's position and orientation when reconstructed from those projections. This method is, in effect, optimized for the choice of ellipse algorithm used; for the purpose of this Note, a minimum bounding ellipse is found based upon the Khachiyan algorithm (see Todd and Yildirim [16]).

In the target's CAD frame, the descriptive spheroid is centered at $\underline{q}_{(t)}$, with its axis of symmetry pointing in the direction $\underline{p}_{(t)}$. A set of n rotation matrices $R_{(i)}$ (i ranging from 1 to n, referring to the ith projection in the series) from the camera coordinate system to the target's CAD coordinate system are then chosen, coupled with vectors $\underline{q}_{(i)}$ to translate the target model to an arbitrary location. Thus, if $\underline{q}_{(t,i)}$ and $\underline{p}_{(t,i)}$ are, respectively, the position and orientation of the spheroid in each simulation, then

$$\underline{q}_{(t,i)} \doteqdot R_{(i)}^T \underline{q}_{(t)} + \underline{q}_{(i)} \tag{7a}$$

$$\underline{p}_{(t,i)} \doteqdot R_{(i)}^T \underline{p}_{(t)} \tag{7b}$$

For each scenario i, given an estimated value of A, B, $\underline{q}_{(t)}$, and $\underline{p}_{(t)}$, the two found pose estimates ($\underline{q}_{(f,j,i)}$ and $\underline{p}_{(f,j,i)}$) of the target can be obtained through spheroid reconstruction (the j index refers to the solution number of that reconstruction). A function $d_{(i)}$ is defined that attempts to encapsulate the error in the reconstruction process for that scenario:

$$d_{(i)} \doteqdot \min_{j=1,2} \left\{ \varpi_1 \frac{(\underline{q}_{(t,i)} - \underline{q}_{(f,j,i)}) \cdot (\underline{q}_{(t,i)} - \underline{q}_{(f,j,i)})}{(\underline{q}_{(i)} \cdot \underline{q}_{(i)})} + \varpi_2 (1 - (\underline{p}_{(t,i)} \cdot \underline{p}_{(f,j,i)})^2) \right\} \tag{8}$$

In Eq. (8), the weighting terms ϖ_i are chosen to reflect any difference in importance between an accurate position estimate and an accurate attitude estimate (in this Note these terms are set to 1). This choice of cost function can be considered geometrically in the following sense: the first term gives a measure of the distance from the CAD model's center to the reconstructed solution of the target; the second term gives a measure of the angular offset from the descriptive spheroid's orientation to the reconstructed orientation.

By defining

$$D \doteqdot \sum_{i=1}^{n} d_{(i)}$$

the minimum of D defines the parameters that describe the descriptive spheroid. The author's definition of D is one of many possible cost functions; however, the determination of this seven-dimensional function's minimum requires only a two-dimensional numerical search method, through appropriate substitution.

The descriptive spheroid is determined through, and therefore influenced by, training data of simulated projections and reconstructions. Considering that the target is projected onto a unit sphere, the effect of the target's position upon the cost function is only due to its radial component. This effect can be attributed to the transition from perspective (finite distance) to orthographic (infinite distance) and the effect on A can be seen in Fig. 3. The solution to the descriptive spheroid's center, orientation and eccentricity $e \doteqdot \sqrt{1-(B/A)^2}$ (for a prolate spheroid) are not affected by the magnitude of $|\underline{q}_{(i)}|$ in the training set. The author's decision to fix the value of $|\underline{q}_{(i)}|$ in the training set means that there is an optimum distance, with regard to the choice of A, at which the accuracy is highest. The optimum distance in this Note is chosen so that the target's projection onto an image plane would take up 15 pixels in length. As the principal concept of this method is to make an initial relative pose estimate when the target comes into visual range, this is arguably the most appropriate optimum range. Fixing the position, the training set is generated by rotating the target to create a uniform distribution of directions for $\underline{p}_{(i)}$, along with further rotations about those axes. Using these methods, a descriptive spheroid for any target model can be obtained.

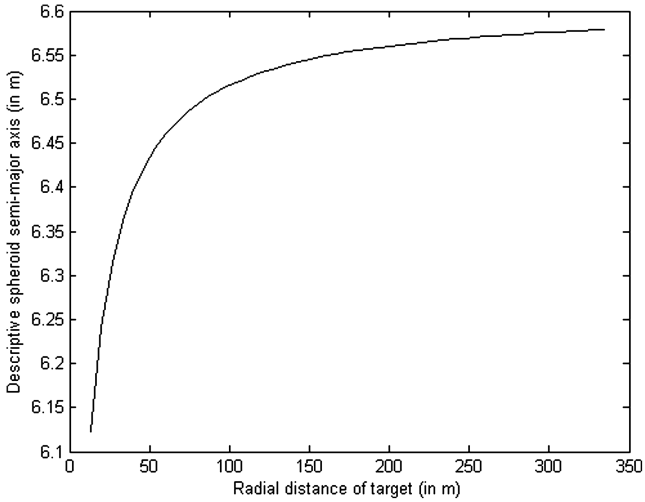


Fig. 3 Effect on A (in meters) of varying the radial distances in the training set when choosing the descriptive spheroid of an elongated target with dimensions 0.5 × 0.5 × 6.5 m^3; modeled around the UoSat-OSCAR 11 satellite (see Bloom [21]).

IV. Intelligent Self-Powered Module: Spheroid Description

In this section, the target CAD model used in this Note is described and its descriptive spheroid is found. The cost function from Eq. (8) is defined with ϖ_1 and ϖ_2 equal to 1. For each reconstruction, however, the solution with the smallest angular discrepancy was chosen.

In this Note the target model considered is the intelligent self-powered module (ISM) (see Fig. 4), which acts as a building block of a larger structure. These modules, designed at the Surrey Space Centre (see Smail and Underwood [17]), respond to the need of building large satellite structures in space from a series of smaller satellites, assembling through magnetic guidance to create a larger structure once in orbit. Two joint ISMs will create a structure that can be used to compare feature-detection methods to the spheroid modeling method, as discussed next.

To describe the characteristics of a single ISM, consider an origin at the center of the ISM and the axes aligned with the electromagnetic flat docking systems (EFDSs) on each facet (the discs). Then, the EFDSs will each lie 0.35 m away from the center. The advantage of considering the ISM target structure in this analysis is that its design is partially engineered to use specific features/markers on the ISM for its pose determination by another spacecraft; the EFDSs are easily extracted from the image and are perfect discs: an example of a degenerate spheroid. Using a technique similar to that described in

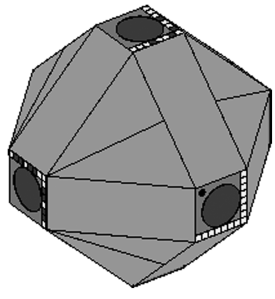
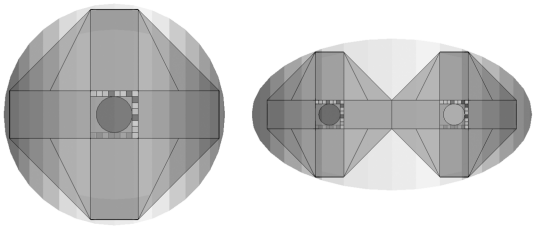


Fig. 4 CAD model of a single ISM.



a) Descriptive Sphere for 1 ISM **b) Descriptive Spheroid for 2 ISMs**

Fig. 5 Descriptive spheroids of single- and double-ISM structures.

this Note, the pose of an ISM structure can be determined from analyzing these markers (see Wokes et al. [18]). In the following section, that method is compared with modeling a two-ISM structure as both a spheroid and a sphere.

The ISM module itself is different from the other targets considered, in that its structure has three orthogonal planes of symmetry. If a single ISM is considered, the process of finding the descriptive spheroid is therefore simpler, as the center of the descriptive spheroid must be at the center of the CAD frame origin and its eccentricity must be 0: a descriptive sphere, with $\underline{q}_{(t)} = (0, 0, 0)$ and a radius of 0.370 m. Without an eccentricity, $\underline{p}_{(t)}$ is undefined; attitude information is not possible.

When there are two ISMs (say, joined along the first axis) a spheroid is obtained; however, by the three planes of symmetry of the structure, the solution for $\underline{q}_{(t)}$ must lie on the intersection of each of them. Likewise, the vector $\underline{p}_{(t)}$ lies along the first axis: the axis on which the ISM centers lie. In this scenario, the descriptive spheroid for two ISMs has a half-length of the axis of symmetry given by 0.788 m and a half-length of the degenerate axes given by 0.425 m. The descriptive spheres/spheroids of single- and double-ISM structures are shown in Fig. 5.

It was noted earlier that the minimum bounding spheroid is not necessarily the spheroid that yields the minimum error when reconstructing from a target's projection. It is observed through Fig. 5 that the descriptive spheroid is, in fact, smaller than the minimum bounding spheroid; there are vertices of the target model that lie outside of the descriptive spheroid. This result is not intuitive, but is understandable with further consideration; if the target was a cube with sides of length l, then its minimum bounding sphere would have radius $l(\sqrt{3}/2)$. If the cube is then translated along the focal axis, so that its center is at $(x, 0, 0)$ and its axes aligned with the camera axes, then its projection onto the image plane would be a square. The minimum bounding circle of its projection would have a radius of $(l\sqrt{2})/(2x - l)$. Meanwhile, the projection of the cube's minimum bounding sphere would give a circle of radius of $(l\sqrt{3})/(\sqrt{4x^2 - 3l^2})$. Taking the ratio of these radii it can be shown that if $(x/l) > \frac{3}{2}$, then the minimum bounding circle is contained in the projection of the minimum bounding sphere. Applying the same reasoning to more complicated target models, it is important to stress that the descriptive spheroid is optimized for reconstructing targets *from their perspective projections*.

The method used by the authors to determine the descriptive spheroid involved parameterizing the dimensions of the spheroid with A and e, rather than A and B. In doing so, it was found numerically that the recovered eccentricity of the descriptive spheroid was given by (for prolate spheroids)

$$e_{\min} \doteqdot \min_{i=1,\ldots,n} \left\{ e \middle| \frac{1}{\sqrt{1-e^2}} \geq \frac{\tan\alpha_i}{\tan\beta_i} \right\} \tag{9}$$

where i spans the length of the training data and α_i and β_i are the semimajor and semiminor arc lengths of the bounding ellipses in the training data. This relation is understood by considering Eq. (4); if the eccentricity were less than $e_{\min}$, there would exist complex reconstructions in the training data, or real reconstructions that first required a modification in the ellipse parameters. Using $e_{\min}$ as the descriptive spheroid's eccentricity, the pose estimation in simulations/practice should never require such modifications.

V. Results

In this section, the theoretical accuracy of the descriptive spheroid is quantified, based upon a perfect nonpixelated image. Following this analysis, an imaging environment is defined and the ISM structure is reconstructed based its their descriptive spheroids and the computational burden is evaluated, relative to more traditional feature-detection methods.

A. Theoretical Accuracy of the Descriptive Spheroid

Once the descriptive spheroid of a target is obtained, the training set can then be used to quantify the magnitude and distribution of reconstruction errors as a result of the heuristic approximation. Because the bounding ellipses found for the training set are based on exact projections of the target's vertices, image noise, lighting and resolution-induced errors are removed from the analysis.

Figures 6 and 7 depict the error in position (given by the magnitude of the vector between the target's reconstructed position and the descriptive spheroid's position) along with the error in orientation (given by the magnitude of the angular difference between the axes of symmetry of the descriptive spheroid and the reconstructed spheroid). The figures showing the distribution of errors in reconstruction have the axes of percentage error against frequency; the percentage error ranges from 0% (a perfect reconstruction) to 100% (a reconstruction with the maximum error obtained for that model) in either position or orientation.

In this Note, only pose estimation of the double-ISM structure is considered for comparison against feature-detection methods. However, by considering the theoretical errors of both the single-ISM and double-ISM structures a trend is observed as the target becomes more accurately described by its conic approximation. The error distribution, from the training data, for the single ISM is position-only: attitude cannot be determined when modeling the target as a sphere; this is shown in Fig. 6. The training data found a maximum error of 0.28 m when the target was 7.38 m away from the camera: a relative error of less than 4%.

For the double-ISM structure, errors in both the position and orientation can be evaluated, also giving a representation of the variation in accuracy as the target's orientation changes. This is shown in Fig. 7. The training data found a maximum error in position of 1.9 m (a mode of 0.6 m) when the target was 14.2 m away from the

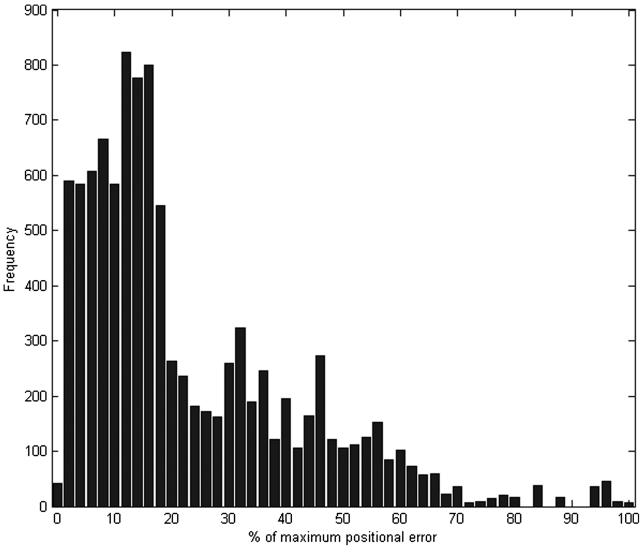
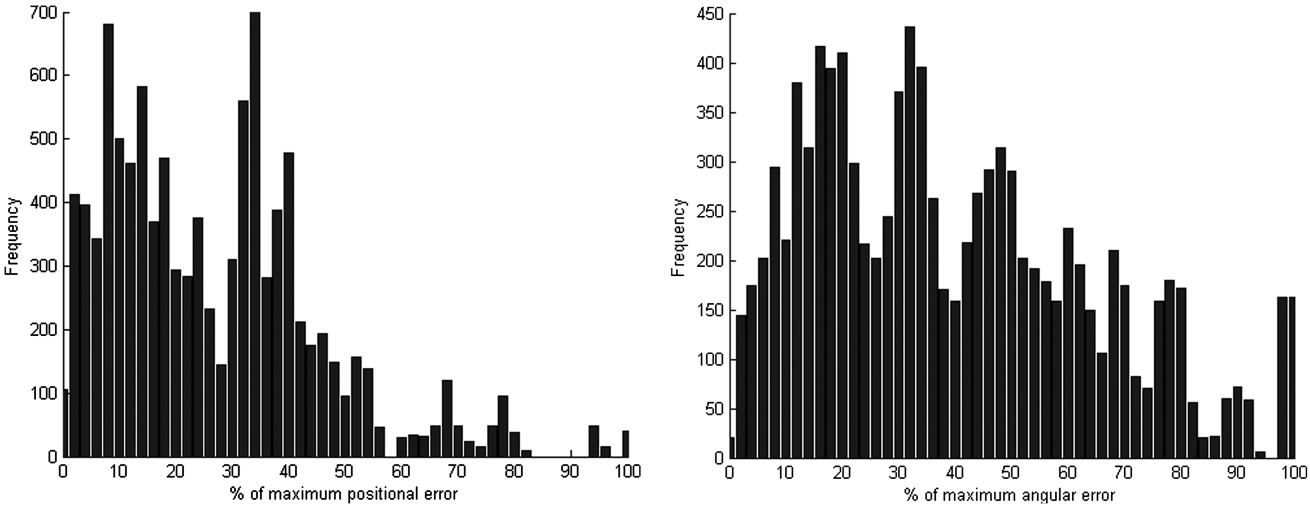


Fig. 6 Distribution of errors (in position) for a single ISM over the training set, as percentages of the largest error found. The maximum error in position was found to be 0.286 m, at a distance of approximately 7.4 m.



a) Error distribution in position **b) Error distribution in orientation**

Fig. 7 Distribution of errors in position and attitude for two ISMs over the training set, as percentages of the largest errors found. The maximum error in position was found to be 1.9 m, and the maximum error in orientation of the axis of symmetry was found to be 41°, at a distance of 14.2 m.

camera: a relative error of 13%. The error in orientation can grow as large as 41° for specific orientations (with a mode of 13°).

B. Defining the Simulated Imaging Environment

For the purposes of displaying and interpreting images, the method used for gathering data and results will be to generate a planar image, find the ellipse on the image plane, and then convert the image plane ellipse into its corresponding image sphere ellipse. Wokes [15] derives a relationship between these ellipses. Considering an image plane, rather than an image sphere, for generating images does not negate the value of projecting onto a sphere and the insight that is gained; the ellipse values of α and β can be checked (and corrected if necessary) to ensure a valid reconstruction from Eqs. (2) and (3). This is not straightforward when relying upon an image plane ellipse alone. Furthermore, every projection conserves particular properties of the space it captures and this is done so through the choice of image surface (or the lenses that precede it). An image sphere preserves the projection of points of equal radial distance and removes the concept of the focal axis. Thus, if a spherical camera were used, the effects of off-axis perspective stretching is removed. For this reason, the target is positioned along the focal axis in all simulations. As our reconstruction equations hold for off-axis positions, moving the target off the focal axis would have no effect on the accuracy of the reconstruction.

POV-Ray (a freeware ray-tracing program[‡]) is used to generate an image using the CAD model, with its position and orientation defined with a call from MATLAB. The image resolution was chosen to be 600 (rows) by 800 (columns) pixels, with an additional background of stars to model noise in the image. The image plane is set at a distance of 1 mm from the focal point of the camera. The top of the image is at 0.5 mm, giving an area of $1 \times \frac{4}{3}$ mm^2. This scaling defines a camera with a $53.130 \times 67.380°$ field of view.

To simulate an arbitrary orientation of the target, the target is rotated 40° about its third axis, 20° about its second axis, and −135° about its first axis (the spheroid's axis of symmetry). The ISM structure is then imaged at various translated distances along the focal axis, from when it fills the entire image (1 m) to when the target is little more than a few pixels (80 m). Simulations at the extremal distances are shown in Fig. 8.

A known disadvantage of passive imagery over active scene interpretation techniques is that the target's appearance is heavily dependent upon the lighting conditions and the sources of light (see Kasai et al. [19]). If the primary light source is behind the target, then the target may disappear completely. Only by simulating a light source from behind the camera itself will all facets of the target object be captured in its image projection. In all of the simulated images, two white light sources are positioned to illuminate the target. These are placed 10 mm along the camera frame's $+y$ and $+z$ axes, respectively, removing the majority of target shadowing (the $+x$ axis is the camera's focal axis).

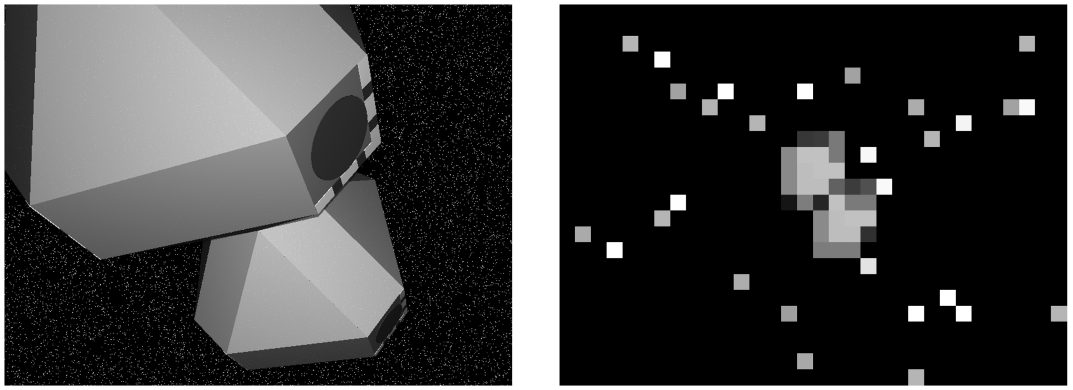
C. Global Versus Feature Modeling

The appearance of the ISM structure was designed with detectable features in mind: the EFDSs are a different material and color from the surrounding hull. Each disc that is extracted from a Portable Network Graphics (PNG) color image can give two possible solutions to an ISM center. The collection of these centers can be analyzed and the pose of the structure can be deduced (see Wokes et al. [18]). Applying this technique, a comparison can be made between the reconstruction accuracy of feature-detection methods and reconstructing the target using a global description. This gives quantitative support that different target descriptions are more appropriate at different radial distances.

Considering a two-ISM target structure, its pose is reconstructed using three different methods: detecting the disc features on its facets and deducing the target's center, describing the structure as a spheroid, or describing the structure as a sphere. In the first case, the features are perfect discs and at lower radial distances pixelation will have little effect on the accuracy of the algorithm. Additionally, this approach also determines the orientation of the structure, even about the axis that connects the two ISMs, modulo 90°. In the second case, a spheroid description will not yield as high an accuracy with lower radial distances; however, as the radial distance increases the effects of pixelation will have less effect on the reconstruction than it will using feature detection. A spheroid description will also give information on the target's orientation, up to its axis of symmetry (see Fig. 7b). There will also be two possible orientations of this axis of symmetry. In the third case, the spheroid is replaced with a sphere and all information about the orientation is lost. However, a single position estimate is reconstructed: an estimate that is less affected by pixelation than that of the spheroid model. Whereas the first method requires color images (or a tinted camera lens) to extract the features by their unique hue, the second and third methods follow a simpler, faster algorithm; a grayscale Portable Gray Map (PGM) image is opened, then the image is thresholded, eroded, or filtered appropriately to detect candidate pixels that capture the target's projection.

Methods of correcting a bounding ellipse have been discussed previously. When two ISMs are modeled with a descriptive sphere, the radius of that descriptive sphere is chosen to be $A = 0.37$ m: the

[‡]Data available online at http://www.povray.org/ [retrieved November 2010].



a) Minimum radial distance (1 m) **b) Maximum radial distance (50 m, zoomed)**

Fig. 8 Illustrations of the considered simulations at the closest and farthest radial distances of the double-ISM structure.

radius for a single ISM. Hence, the ellipse correction method used on an image for reconstructing the descriptive sphere is to reduce the semimajor axis of the image sphere ellipse to match the semiminor axis. This implies that if the ISM structure were oriented perpendicular to the camera's focal axis, then the descriptive sphere would give as accurate a result as the descriptive spheroid.

The positional errors of the three target reconstruction methods are given in Fig. 9. As the reconstructed orientation varies in detail, depending on which reconstruction method used, such comparisons are omitted. The increase in relative positional error highlights the effects of pixelation at higher distances.

Figure 9a shows that for radial distances up to 20 m, the feature-detection method gives an almost consistent higher accuracy in the position determination. After 20 m, the accuracy deteriorates rapidly until, after 26 m, the features cannot be accurately extracted from the image noise and the method fails. In Fig. 9b, the descriptive spheroid gives a better position accuracy than the descriptive sphere, but the difference between these errors becomes insignificant at greater distances. The fundamental advantage, still at 80 m, is the ability to determine an attitude estimate of the target by using a descriptive spheroid.

The radial distances considered for this simulation ranged from 1 m to 80 m. CMOS cameras range in resolution with current commercial high-resolution color cameras (for example, the FCi4-14000 model from C-Cam Technologies) yielding 14 megapixels (Iwane et al. [20]): roughly five times greater detail than the 800×600 pixel images considered here. By a simple scaling, this means that the accuracy found in simulations can be assumed for five times the calculated distance. Therefore, for the ISM structure using a wide-angle camera, feature-detection techniques may fail at 100 m, whereas a global model can still give both position and limited attitude estimates up to 400 m away.

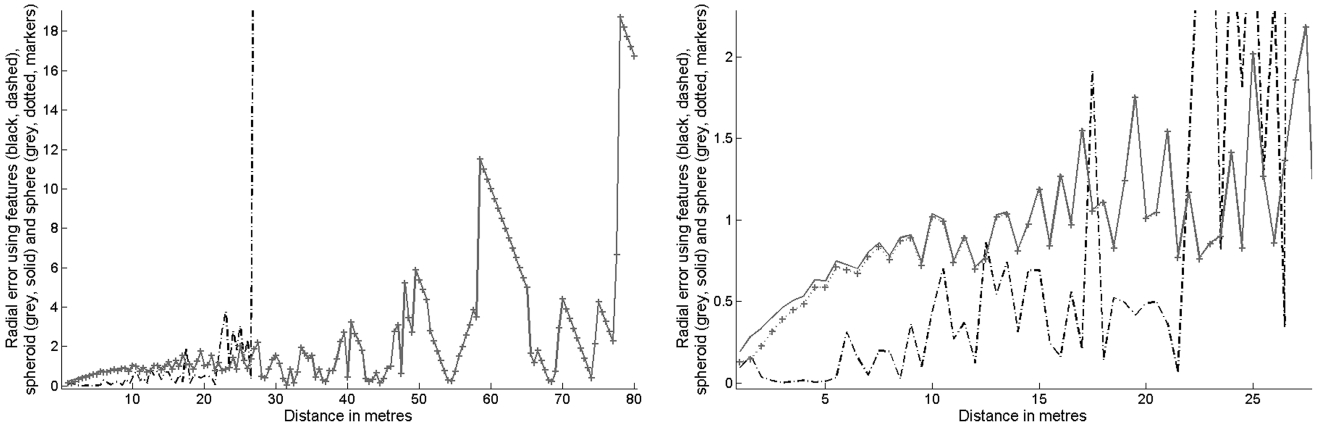
D. Evaluation of Computational Burden

To address the question of computational burden, the simulations were timed using the profiler in MATLAB. The laptop used was running on a Windows Vista operating system, with a 1.83 GHz processor and 3 GB of RAM. However, in order to remove the processor dependence when these algorithms are evaluated, the runtimes are also given in terms of the ratio Ξ, defined as

$$\Xi \doteq \frac{T_1}{T_2} \tag{10}$$

where T_1 is the concatenation of the image processing, ellipse bounding and target reconstruction runtime, and T_2 is the runtime for opening/reading the image (the `Imread` function). The numerator of Ξ encapsulates the time for analyzing the image, extracting the target from the background, finding the bounding ellipse of the target's pixels, and reconstructing the original target's relative position. The denominator of Ξ is the time for opening the image file in MATLAB: a process that any image-based pose estimation algorithm would have to complete. Ξ gives a measure of the calculation speed obtained that is independent of the workstation used. Thus, if $\Xi = 1$, then the time it takes to extract the relative pose estimate from the image is equal to the time it takes to open that image. A 0 value of Ξ means that the reconstruction process takes no time at all.

Using unoptimized MATLAB code, the feature-based pose determination algorithm can have a runtime from 1.7 s (corresponding to $\Xi = 1.32$) to 8.3 s (corresponding to $\Xi = 10.36$). The increase in computation time is caused when only single facets are detected and a more complex algorithm is used. The greatest computational burden occurs with the transformation of the image from red–green–blue arrays to hue–saturation-value arrays: an



a) Full radial distance against positional error **b) Magnified radial distance against positional error**

Fig. 9 Error comparison (in meters) between feature detection (black dashed line), spheroid modeling (gray solid line) and sphere modeling (gray dotted markers).

unavoidable aspect of feature detection if the features are distinguished by their color. Conversely, reconstructing the ISM structure's pose (or position) from an image without analyzing its features has a runtime of 0.21 s. This corresponds to a ratio Ξ of 0.74.

Whether these calculation times are suitably fast for autonomous rendezvous is determined by the predefined closing rates between the target and the chaser satellite, which is scenario-specific. Cropp's pose determination algorithms (mentioned in Sec. I) had a runtime of 60 s on a 400 MHz processor (see Cropp [11]). This corresponds (approximately) to a runtime of 13.3 s on the laptop used for determining the previous runtimes in this article and is dependent on a heuristic input. Comparing these times, the heuristic input takes 0.7% of the calculation time for accurately determining the relative pose through edge-line correspondence.

VI. Conclusions

This Note introduces the concept of obtaining a fast pose estimate of a target by modeling that target as a spheroid. For a general target, feature-detection methods can give reliable relative pose estimates up to a given distance. Comparing these accuracies with a target's reconstruction when (suitably) modeled as a spheroid or sphere, there is a radial boundary beyond which it is better to model the target globally. The smaller the features used for detecting and tracking, relative to the target's global dimensions, the greater the distance for which heuristic modeling methods will be applicable when the former method will not.

The advantages of modeling the target object as a spheroid are that a fast pose estimate is obtained with low computational cost and is not dependent on a complex analysis of the image received. There is no assumption of communication between the target and chaser satellite or any distinct features or markings on the target that are required other than obtaining an estimate to the overall dimensions. More critically, this Note experimentally validates an intuitive concept: at larger distances, a global description of a target gives better pose estimation technique than attempting to identify smaller features, whose accuracy is affected to a greater extent by pixelation.

References

[1] Besl, P. J., and Jain, R. C., "Three Dimensional Object Recognition," *ACM Computing Surveys*, Vol. 17, No. 1, 1985, pp. 75–145. doi:10.1145/4078.4081

[2] Lowe, D. G., "Three Dimensional Object Recognition from Single Two-Dimensional Images," *Artificial Intelligence*, Vol. 31, No. 3, March 1987, pp. 355–395. doi:10.1016/0004-3702(87)90070-1

[3] Ruel, S., English, C., Anctil, M., and Church, P., "LASSO: Real-Time Pose Estimation from 3D Data for Autonomous Satellite Servicing," *8th International Symposium on Artificial Intelligence, Robotics and Automation in Space*, edited by B. Battrick, ESA, 2005, pp. 78–85.

[4] Davidovitz, A., and Shinar, J., "Two-Target Game Model of an Air Combat with Fire-and-Forget All-Aspect Missiles," *Journal of Optimization Theory and Applications*, Vol. 63, No. 2, Nov. 1989, pp. 133–165. doi:10.1007/BF00939571

[5] Hmam, H., and Kim, J., "Optimal Non-Iterative Pose Estimation via Convex Relaxation," *Image and Vision Computing*, Vol. 28, No. 11, Nov. 2010, pp. 1515–1523. doi:10.1016/j.imavis.2010.03.005

[6] Agarwal, S., and Chaudhuri, S., "Determination of Aircraft Orientation for a Vision-Based System Using Artificial Neural Networks," *Journal of Mathematical Imaging and Vision*, Vol. 8, No. 3, May 1998, pp. 255–269. doi:10.1023/A:1008226702069

[7] Beaulieu, M., "Launch Detection Satellite System Engineering Error Analysis," M.S. Thesis, Naval Postgraduate School, Monterey, CA, March 1996.

[8] Stieber, M. E., McKay, M., Vukovich, G., and Petriu, E., "Vision-Based Sensing and Control for Space Robotics Applications," *IEEE Transactions on Instrumentation and Measurement*, Vol. 48, No. 4, Aug. 1999, pp. 807–812. doi:10.1109/19.779178

[9] Shakunaga, T., "An Object Pose Estimation Using a Single Camera," *Proceedings of the 1992 lEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 2, IEEE, Piscataway, NJ, July 1992, pp. 1053–1060.

[10] Lowe, D. G., "Fitting Parameterized Three-Dimensional Models to Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 5, 1991, pp. 441–450. doi:10.1109/34.134043

[11] Cropp, A., "Pose Estimation and Relative Orbit Determination of a Nearby Target Microsatellite using Passive Imagery," Ph.D. Thesis, Electronic Engineering, Information Technology and Mathematics, Surrey Space Centre, Univ. of Surrey, Guildford, England, U.K., April 2001.

[12] Fischler, M. A., and Bolles, R. C., "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Communications of the ACM*, Vol. 24, No. 6, June 1981, pp. 381–395. doi:10.1145/358669.358692

[13] Eberly, D. H., *3D Game Engine Design: A Practical Approach to Real-Time Computer Graphics* , 1st ed., The Morgan Kaufmann Series in Interactive 3D Technology, Morgan Kaufmann, New York, Sept. 2000.

[14] Wokes, D. S., and Palmer, P. L., "Perspective Reconstruction of a Spheroid from an Image Plane Ellipse," *International Journal of Computer Vision*, Vol. 90, No. 3, 2010, pp. 369–379. doi:10.1007/s11263-010-0368-0

[15] Wokes, D. S., "Autonomous Pose Estimation of a Passive Target," Ph.D. Thesis, Surrey Space Centre, Univ. of Surrey, Guildford, England, U.K., Jan. 2010.

[16] Todd, M. J., and Yildirim, E. A., "On Khachiyan'S Algorithm for the Computation of Minimum-Volume Enclosing Ellipsoids," *Discrete Applied Mathematics*, Vol. 155, No. 13, 2007, pp. 1731–1744. doi:10.1016/j.dam.2007.02.013

[17] Smail, S., and Underwood, C., "Electromagnetic Flat Docking System for In-Orbit Self-Assembly of Small Spacecraft," *AAS/AIAA Space Flight Mechanics Meeting*, American Astronautical Society, Springfield, VA, Feb. 2009.

[18] Wokes, D. S., Smail, S., Palmer, P. L., and Underwood, C. I., "Pose Estimation for In-Orbit Self-Assembly of Intelligent Self-Powered Modules," AIAA Guidance, Navigation, and Control Conference and Exhibit, Chicago, AIAA Paper 2009-6291, 2009.

[19] Kasai, T., Oda, M., and Suzuki, T., "Results of the ETS-7 Mission—Rendezvous Docking and Space Robotics Experiments," *Fifth International Symposium on Artificial Intelligence*, ESA, Noordwijk, The Netherlands, June 1999, pp. 299–306.

[20] Iwane, M., Matsuda, T., Sugai, T., Tazoe, K., Okagawa, T., Ono, T., Watanabe, T., Ogawa, K., Takahashi, H., and Inoue, S., "52 Mega-Pixel APS-H-Size CMOS Image Sensor for Super High Resolution Image Capturing," *International Image Sensor Workshop*, Ogunquit Maine, June 2007, pp. 295–298.

[21] Bloom, J., "A Profile of the UoSat-OSCAR 11 Satellite," *The ARRL Satellite Anthology*, American Radio Relay League, Newington, CT, 1988, pp. 52–53.