

## Research Article

# Pose Estimation of Uncooperative Unknown Space Objects from a Single Image

Xiaoyuan Ren, Libing Jiang, and Zhuang Wang 

*Automatic Target Recognition Laboratory, National University of Defense Technology, 410073, China*

Correspondence should be addressed to Zhuang Wang; [zhuang\\_wang@sina.com](mailto:zhuang_wang@sina.com)

Received 28 February 2020; Accepted 6 May 2020; Published 18 July 2020

Academic Editor: Hikmat Asadov

Copyright © 2020 Xiaoyuan Ren et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Estimating the 3D pose of the space object from a single image is an important but challenging work. Most of the existing methods estimate the 3D pose of known space objects and assume that the detailed geometry of a specific object is known. These methods are not available for unknown objects without the known geometry of the object. In contrast to previous works, this paper devotes to estimate the 3D pose of the unknown space object from a single image. Our method estimates not only the pose but also the shape of the unknown object from a single image. In this paper, a hierarchical shape model is proposed to represent the prior structure information of typical space objects. On this basis, the parameters of the pose and shape are estimated simultaneously for unknown space objects. Experimental results demonstrate the effectiveness of our method to estimate the 3D pose and infer the geometry of unknown typical space objects from a single image. Moreover, experimental results show the advantage of our approach over the methods relying on the known geometry of the object.

## 1. Introduction

The pose estimation of uncooperative space objects is a key technology for many space missions such as target recognition and on-orbit service [1, 2]. In contrast to other state-of-the-art systems, a monocular camera ensures pose estimation under low power and low hardware complexity [3]. However, it is sometimes hard for the monocular camera to provide clear image sequence limited by the observation condition. Therefore, it is necessary to study the pose estimate of the space object from a single image.

3D pose estimation from a single image is an important but challenging task. For space objects, most of the existing methods rely on the known geometry of the object, which can be broadly classified into model-based and template-based methods. For model-based methods, the 3D pose is determined by an iterative algorithm to minimize a certain fit error between the features detected in the input image and the corresponding features of the known reference 3D model [4–8]. For the template-based methods, the input image is searched for specific features that can be matched to an assigned template, and the 3D pose is obtained from

the best matching with the prebuilt template library of the object [9–13]. Currently, with the development of deep learning technology, a pose determination method based on convolutional neural networks (CNN) is provided in [14]. All of the above methods rely on the known geometry of the object and work on the specific object.

However, the detailed geometry of the object cannot always be known in some scenarios, such as for the unknown space objects. Even for the object with known geometry, there might be difference between the known reference model and the actual object. This is because the structure of the space object may change in different working modes, such as the rotation of solar sails. Compared with the known object, the pose estimation of the unknown object is more difficult. Methods relying on the known object geometry are not available for the unknown objects. For the unknown object, not only the pose but also the shape should be estimated. A few works exist in the literature on pose estimation of unknown space objects. In [15, 16], the feature points of the unknown space object are stored during continuous observation, and the pose is determined through the matching of feature points. However, the stable and continuous observation is always

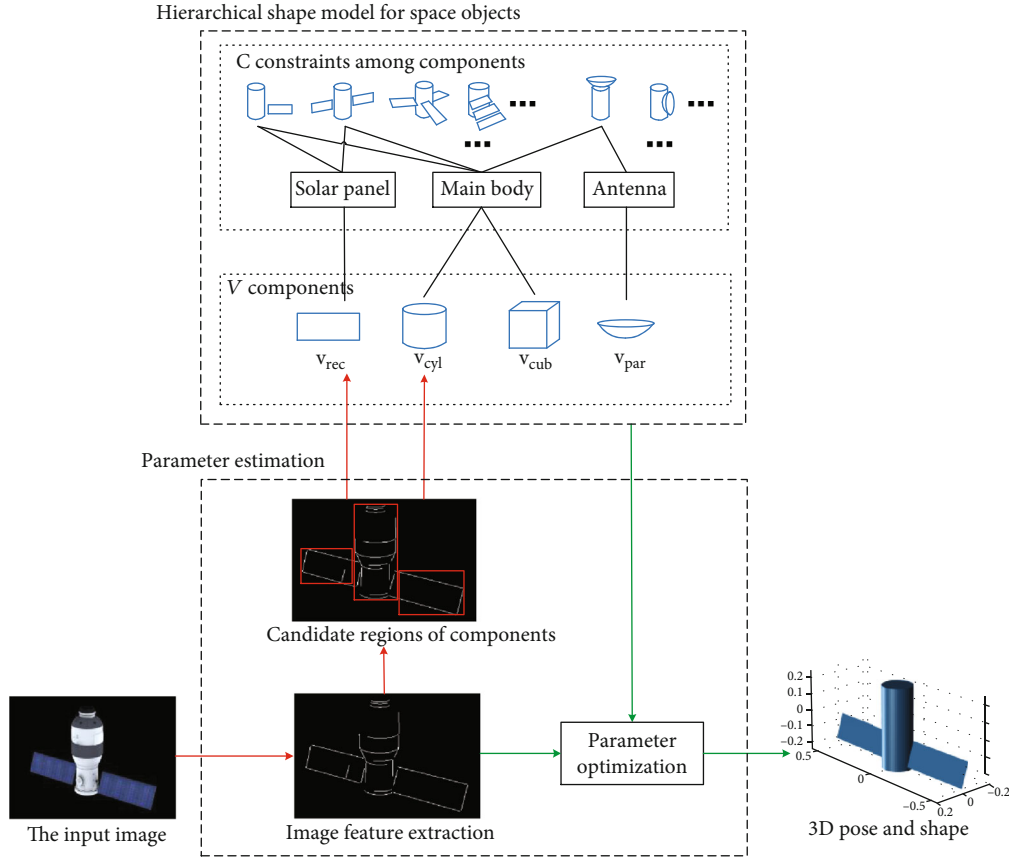


FIGURE 1: The framework of our method. The upper part is the prebuilt hierarchical shape model. The lower part shows the estimation of the shape and pose from the input image. The red line represents the process of activating the hierarchical shape model, and the green line represents the process of parameter estimation in support of the hierarchical shape model.

hard to provide due to the limitation of observation conditions. To our knowledge, there is no effective method at present to estimate the 3D pose of the unknown space object from a single image.

This paper devotes to estimating the 3D pose of uncooperative unknown space objects from a single image. Although this is an ill-posed problem, it can be made possible by introducing the prior information of the object structure. Actually, there has been some researches on this problem in computer vision. Generally, a prior shape model is built based on the structural similarity of objects in a specific category, and the pose estimation of the unknown object is reduced to a 3D-to-2D shape fitting problem where the parameters of shape and pose are estimated simultaneously [17–19]. However, this solution is not practicable for space objects. Compared with common objects such as cars and chairs, there are large structural differences among different space objects. For example, the number of the solar panels, the shape of the main body, and the size of the antenna vary from different space objects. As a result, the structures of different space objects are hard to be represented by a unified model, which causes difficulty to introduce the prior structural information to parameter estimation of the unknown space object.

This paper presents an efficient method to estimate the 3D pose of unknown space objects from a single image. Although space objects have large structural variability, there

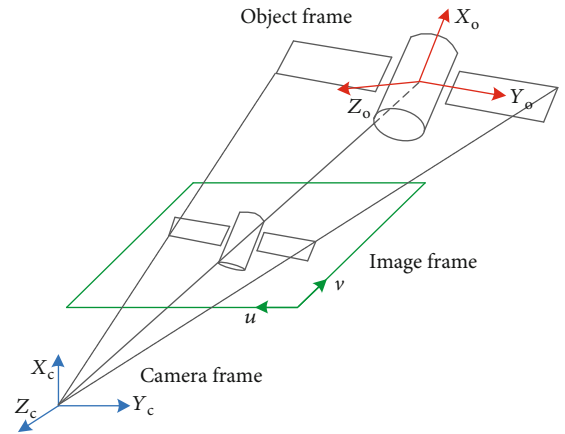


FIGURE 2: Reference frames.

are plenty of constraints among the components of typical space objects. For example, the solar panels of one object always have the same size and spatially symmetric disposition. These constraints have potential value to estimate the pose and shape for unknown space objects. From this, we propose a hierarchical shape model of space objects to describe the prior information of the object geometry, which represents the constraints among components of objects in

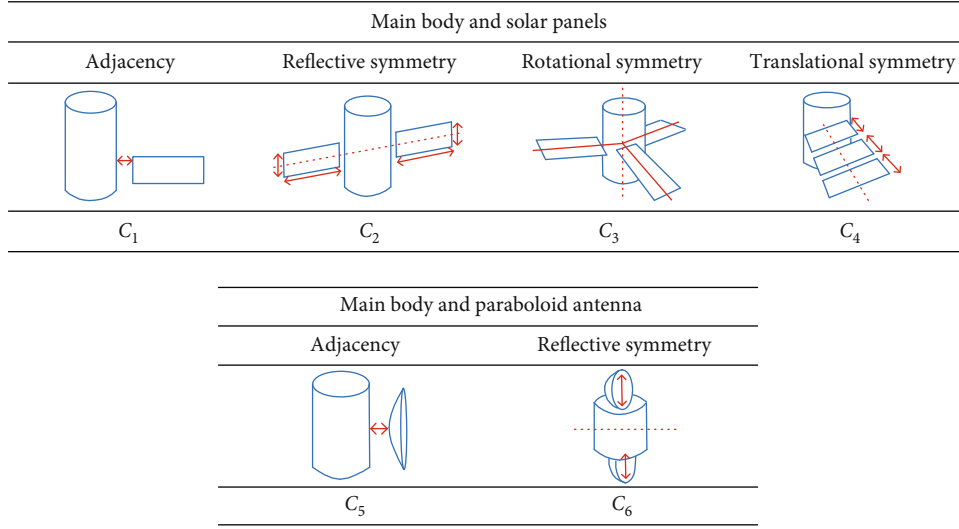


FIGURE 3: The constraints among the components of typical space objects considered in this paper.

the form of probability distribution. Compared with traditional prior shape models in computer vision such as wire-frame [17–19] and mesh [20], the hierarchical shape model proposed is able to describe objects with large structural variability. Therefore, with the support of the hierarchical shape model proposed, the pose estimation of unknown space objects can be conducted from a single image.

The main contribution of this paper is a pose estimation method of unknown space objects from a single image. Specifically, we establish a hierarchical shape model to describe the prior structures of typical space objects. On this basis, the parameter estimation method of unknown space objects is presented.

## 2. Overview

Our method consists of two aspects. Firstly, the hierarchical shape model of typical space objects is established previously. Secondly, the pose and shape of the unknown object are estimated simultaneously from a single image in support of the hierarchical shape model. The framework of our method is shown in Figure 1.

Section 3 illustrates the building of the hierarchical shape model of space objects. Section 4 introduces the estimation of pose and shape.

## 3. Prebuilt Hierarchical Shape Model

This paper builds the hierarchical shape model previously to describe the prior structures of typical space objects. The prebuilt hierarchical shape model is established from the structural laws of space objects, which is defined as  $\partial$ .  $\partial$  is a 2-tuple

$$\partial = \langle V, C \rangle, \quad (1)$$

where  $V$  represents the types of object components and  $C$  indicates the constraints among object components. Each

TABLE 1: Detailed constraints of reflective symmetry.

	$f(\cdot)$	$p$
$c_1$	$ \theta_1 - \theta_2 $	$e^{-\rho_1 \cdot  \theta_1 - \theta_2 }$
$c_2$	$ \mathbf{S}_1 - \mathbf{S}_2 $	$e^{-\rho_2 \cdot  \mathbf{S}_1 - \mathbf{S}_2 }$
$c_3$	$\text{Dis}(\text{axis}_1, \text{axis}_2)$	$e^{-\rho_3 \cdot \text{Dis}(\text{axis}_1, \text{axis}_2)}$

node  $v_i \in V$  is a type of basic shapes corresponding to one type of object components. Each node  $C_i \in C$  is a set of constraints of specific components. The detailed illustration of  $V$  and  $C$  is as follows.

**3.1. Object Components.**  $V$  represents the components of space objects. The components of the space object always can be simplified as the basic shapes. In this paper,  $V$  consists of four types of basic shapes: rectangle, cylinder, cube, and paraboloid, shown as follows.

$$V = \{v^{\text{rec}}, v^{\text{cyl}}, v^{\text{cub}}, v^{\text{par}}\}, \quad (2)$$

where rectangle is used to represent the solar panels, cylinder and cube can be seen as the main body of the space object, and paraboloid indicates the antenna. Admittedly, these basic shapes certainly cannot represent all the components of space objects. Nevertheless, these basic shapes share the general characters of space objects and have the ability to describe the main structure of typical space objects. Besides, the types of basic shapes in our model are extensible.  $v$  has a set of attributes  $\phi(v)$  that describe its position, size, and pose, respectively, as below.

$$\phi(v) = \{\mathbf{X}, \mathbf{S}, \boldsymbol{\theta}\}. \quad (3)$$

$\mathbf{X}$  indicates the position of the component in the object frame.  $\mathbf{S}$  represent the size of the component in the object

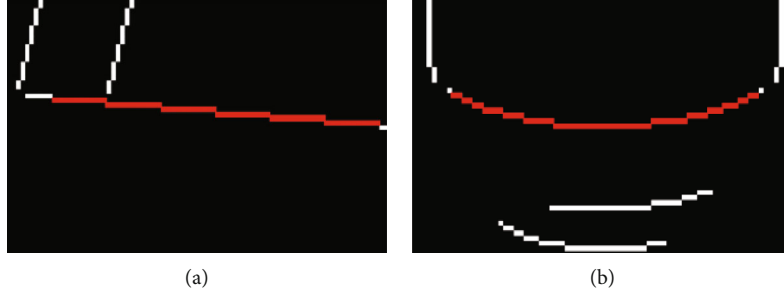


FIGURE 4: Extraction of image features. The line and arc detected are marked in red, respectively, in (a) and (b).

frame.  $\theta$  is the rotation angle between the object frame and the camera frame. Figure 2 shows the reference frames of our method.

**3.2. Constraints among Object Components.**  $C = \{C_1, C_2, \dots, C_{N(C)}\}$  represents the constraints among the components of space objects. The constraints considered in this paper are shown in Figure 3, which exist commonly in typical space objects.

Each  $C_i$  consists of a set of  $c$  defined as

$$c = \{f(\{\phi(v_i)\}_k), p\}, \quad (4)$$

where  $f(\cdot)$  is the function over the attributes of corresponding components of the object and  $p$  is the probability distribution of the responses of  $f(\cdot)$ .

To explain detailedly, Table 1 elaborates  $c_i$  of  $C_2$ , which indicates the relation of reflective symmetry, where  $axis_i$  is the major axis of the object component.  $Dis(axis_1, axis_2)$  indicates the distance between axes of two components.  $c_1$ ,  $c_2$ , and  $c_3$  measure the parallelism, conformance of size, and coaxiality between two components, respectively. Similarly, each  $C_i$  can be represented in this way.

#### 4. Pose and Shape Estimation

On the basis of the prebuilt hierarchical shape model, the pose and shape of the unknown space object can be estimated from the input image. The process of estimation is depicted in the lower part in Figure 1.

Firstly, image features are extracted from the input image, and the candidate regions of object components are determined based on the image features, which are detailed in Section 4.1. Secondly, according to the image features and regions of components, the pose and shape are estimated under the support of the prebuilt hierarchical shape model of space objects, which are detailed in Section 4.2. Besides, this parameter estimation is a complicated optimization problem, and the strategy of optimization is given in Section 4.3.

**4.1. Image Processing.** For the input image, the first step is the feature extraction. In our method, the features extracted from the input image are lines and arcs. This is because the components of space object always can be treated as the simple basic

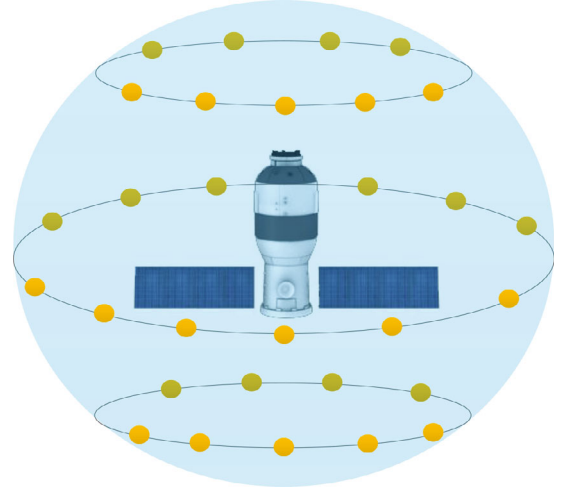


FIGURE 5: Illustration of the pose space discretization using the sphere with uniformly distributed camera locations. Every orange circle in the figure indicates the location of one camera.

shapes such as cylinders and cube, which reflect significant lines and arcs in the image.

The feature extraction method in this paper is inspired by [21]. In [21], connecting pixels are extracted from the gradient magnitudes of the image, and then, lines are detected by aggregation of the connecting pixels. Similar to [21], lines and arcs are extracted, respectively, from the input image by making corresponding aggregation rules in this paper. Figure 4 shows the extraction of image features. The image feature detected in our method is defined as  $L$ . Experiments in Section 5.3 verify the advantage of image feature used in our method.

After the feature extraction, the candidate regions of object components are determined based on the lines and arcs detected. For example, the candidate region of the solar panel contains the long parallel lines, the candidate region of a cylindrical main body contains the parallel lines and arcs, the candidate region of a cuboid main body contains several sets of the parallel lines, and the candidate region of a paraboloid antenna contains arcs.

**4.2. Estimation of Pose and Shape.** Under the hierarchical shape model, an arbitrary object can be represented as  $T = \langle V_T, C_T \rangle$  which is an instance of  $\partial$ , where  $V_T \subseteq V$  and  $C_T$

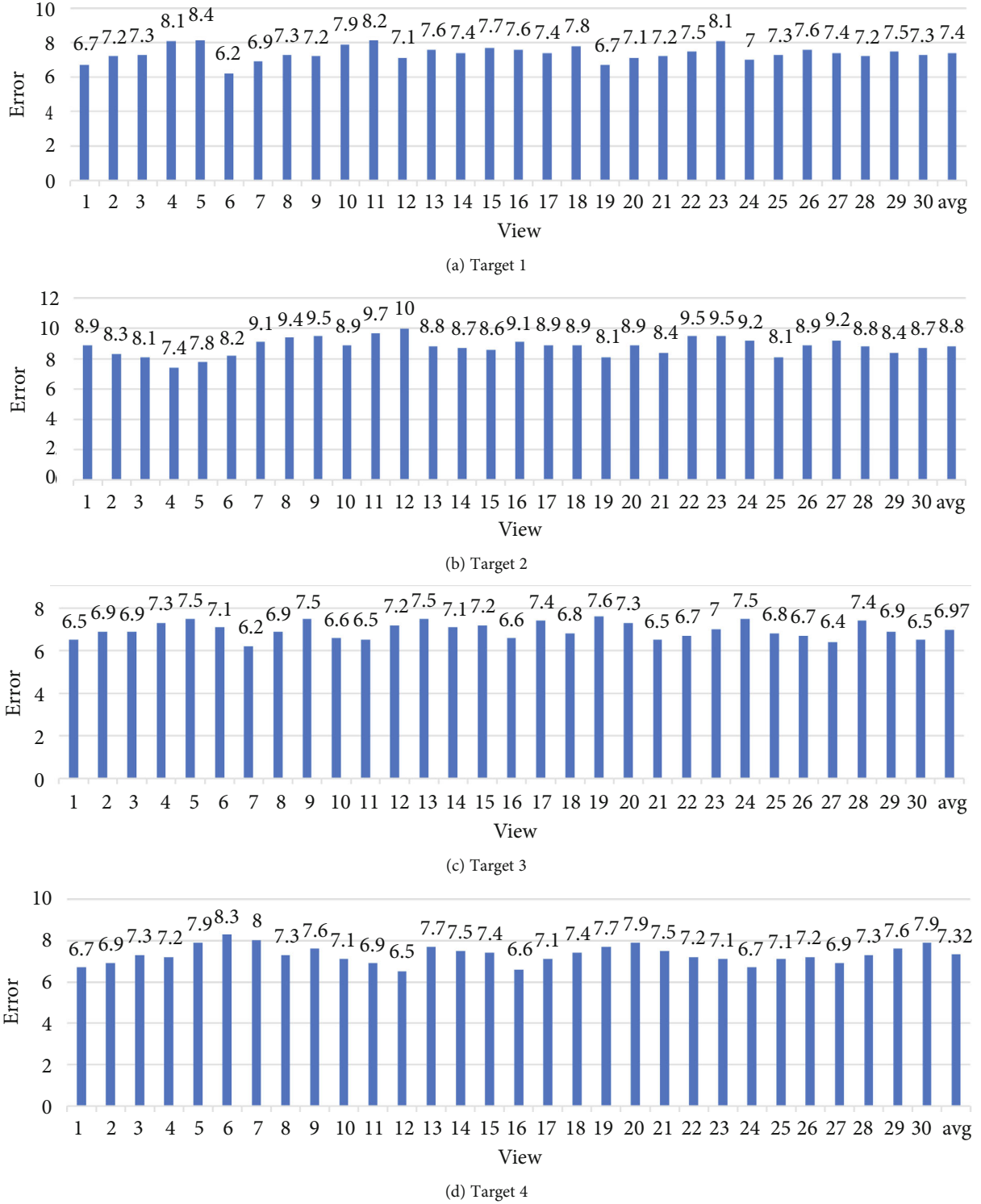


FIGURE 6: Quantitative results of the method proposed. The results of four objects are shown from (a) to (d), respectively. Every object has 30 test images from different views. The performance of the method is evaluated by (10).

$\subseteq C$ . In the support of the hierarchical shape model, estimation of the pose and shape is converted to determine  $T^*$  corresponding to the input image.

As a single image is input, the nodes of the hierarchical shape model are activated.  $T^*$  is obtained by maximum a posteriori probability  $P(T|I)$  as follows:

$$T^* = \arg \max_T P(T|I) = \arg \max_T (P(T) \cdot P(I|T)), \quad (5)$$

where  $P(T)$  is the probability distribution of activated nodes in the hierarchical shape model, which is the product of  $P(C_T)$  and  $P(V_T|C_T)$ . The conditional probability

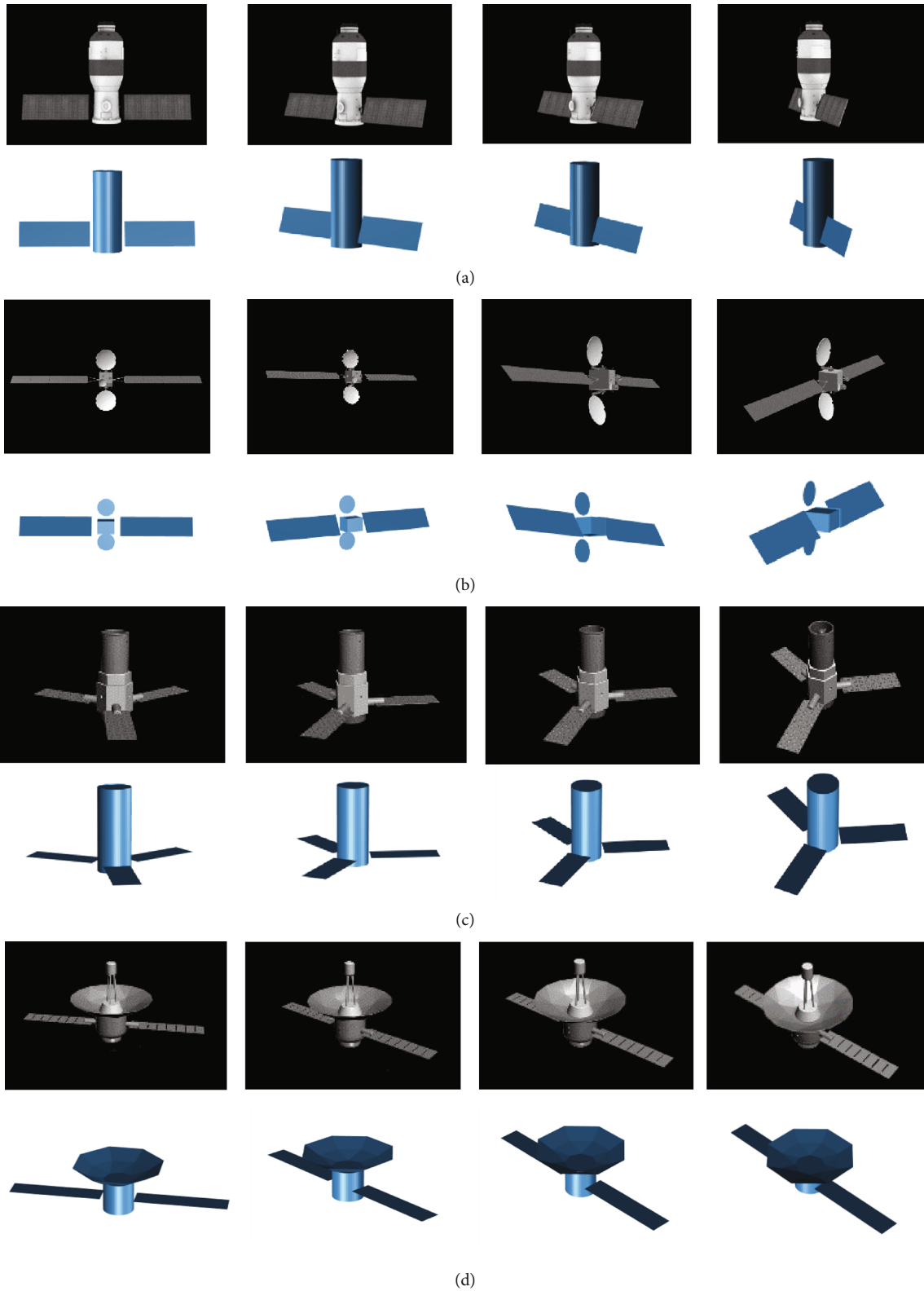


FIGURE 7: Part of the qualitative results of pose estimation of unknown space objects. The bottom of each test image is the object shape inferred by our method under the estimated pose. Four objects are arranged in sequence from (a) to (d).



$P(I | T)$  measures the matching degree of the inferred object and the input image, which is equal to  $P(I | V_T)$ . Consequently, (5) can be represented as

$$\langle V_T^*, C_T^* \rangle = \arg \max_{V_T, C_T} (P(C_T) \cdot P(V_T | C_T) \cdot P(I | V_T)). \quad (6)$$

The first item  $P(C_T)$  in (6) indicates the probability of each type of constraint, which is regarded as a constant in this paper.

The second item  $P(V_T | C_T)$  in (6) can be expressed as follows due to the definition of constraints among the components in (4).

$$P(V_T | C_T) = \prod_{i=1}^{N_{C_T}} P(\{v_k\} | c_i) = \prod_{i=1}^{N_{C_T}} p_i. \quad (7)$$

The last item  $P(I | V_T)$  in (6) measures the matching degree of the inferred object and the input image. For the input image, lines and arcs are detected as image features.  $P(I | V_T)$  is defined as the matching degree between the distance map of image features  $\text{dismap}(L)$  and edge of projection of the inferred object  $V_T^{\text{edge}}$ .

$$P(I | V_T) = e^{-\lambda \cdot \phi_{\text{dist}}(\text{dismap}(L), V_T^{\text{edge}})}, \quad (8)$$

where  $\phi_{\text{dist}}(\cdot)$  measures the difference of the distance map and edge. At this point, the constraints among components of space objects are represented as the probability distribution.

**4.3. Optimization Strategy.** Substituting (7) and (8) into (6), we have

$$\{V_T^*, C_T^*\} = \arg \max_{V_T, C_T} \left\{ \prod_{i=1}^{N_{C_T}} p_i \cdot e^{-\lambda \cdot \phi_{\text{dist}}(\text{dismap}(L), V_T^{\text{edge}})} \right\}. \quad (9)$$

To solve this complicated optimization problem, the model inference strategy is given as follows: (1) The pose of the main body of the space object is estimated firstly as the initial value of other components. This is because the main body can be always seen as a cylinder or cuboid, which has significant difference of image features in different poses. (2) The pruning is used to remove incompatible proposals in the inference of the hierarchical shape model. (3) The particle swarm optimization [22] is used in model inference.

## 5. Experiment

The experiments consist of three parts. In Section 5.1, the pose estimation method proposed is evaluated on four types of unknown space objects. The purpose of this part is to verify the ability of our method to estimate the 3D pose of unknown space objects. In Section 5.2, the accuracy of pose estimation on the object with a changed structure is compared between our method and the mainstream method which relies on the known geometry of the object. This part

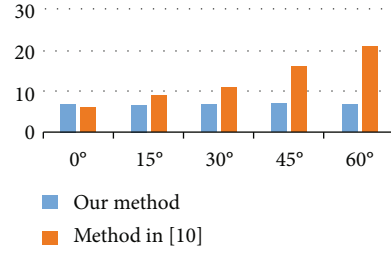


FIGURE 8: The performance comparison of pose estimation between the method in [10] and our method on the object with a changed structure. The accuracy is measured by the median error of different views for each object.

reflects the advantage of our method over the approaches only for the known objects. In Section 5.3, the performance of the image feature extracted mentioned in Section 4.1 is compared with other common features based on the framework of our method. This part is to prove the validity of the image features extracted in our method for space objects.

**5.1. Pose Estimation of Unknown Objects.** The performance of our method is evaluated on four types of unknown space objects from a single image. The experiment data is produced by the computer simulation. For each object, 30 test images are sampled from different views shown in Figure 5.

The error of pose estimation  $E_R$  between the predicted rotation vector  $R_{\text{pred}}$  and the ground truth rotation vector  $R_{\text{GT}}$  is measured as

$$E_R = \frac{\left\| \logm(R_{\text{pred}}^T R_{\text{GT}}) \right\|_F}{\sqrt{2}}. \quad (10)$$

The quantitative results of four objects are shown sequentially in Figure 6.

The qualitative results are shown in Figure 7.

In our method, the shape inference and pose estimation are conducted simultaneously. The object shape inferred is represented as the combination of basis shapes. Experiments show the effectiveness of our method to estimate the 3D pose of unknown space objects.

**5.2. Pose Estimation of the Object with Structure Changed.** Most of the existing methods rely on the known geometry of space object. However, there will always be a difference between the known reference model and the actual object. For the pose estimation of the object with a changed structure, the performance of our method is compared with the method in [10]. The pose estimation in [10] is based on the matching with the template library reflecting the object features in different poses, which is the representative way at present. The experimental object of this part is the first object in Section 3.1. The angle deviation of the solar panels and main body is 0°, 15°, 30°, 45°, and 60°, respectively. The experimental result is shown in Figure 8.

It can be seen that with the deviation of angle increases, the accuracy of the method in [10] decreases significantly. In comparison, our method performs robustly in the case

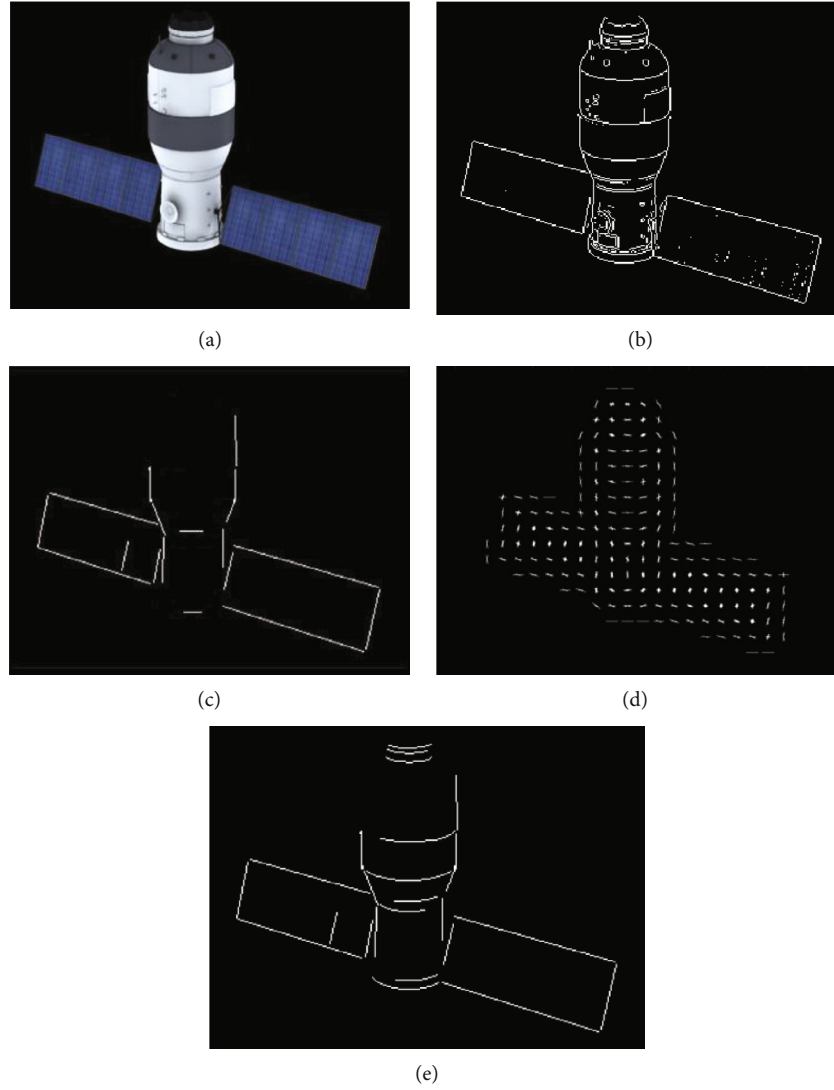


FIGURE 9: Different image features: (a) the original image, (b) the edge of the image, (c) the line feature, (d) the HOG feature, and (e) the image feature adopted in our method.

of change of object structure. The result shows the limitation of the methods relying known object geometry. The experimental result shows the necessity to estimate the pose of the unknown objects.

**5.3. Comparison of Image Features.** In our method, the features extracted from the input image are lines and arcs. In order to verify the validity of feature extracted in this paper, the experiment is designed as follows. Based on the framework of our method, the performance of image feature extracted in our method is compared with several common features including edge [23], line [21], and HOG [24]. These features are displayed in Figure 9.

The quantitative comparison between the image feature in our method and other features is shown in Figure 10.

From the results, the performance of the image feature extracted in our method is better than that of other features. This is because the image features extracted in our method reflect the characteristics of space objects well.

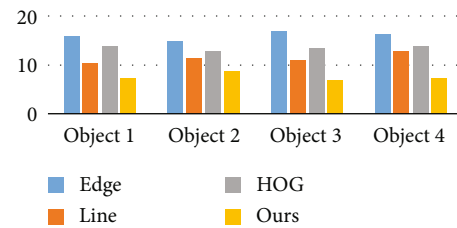


FIGURE 10: The performance comparison of the image features extracted in our method with several common features. The average error is measured on four objects mentioned in Section 5.1.

## 6. Conclusion

This paper provides a pose estimation method for uncooperative unknown space objects from a single image. In this paper, a hierarchical shape model is established to represent the prior structure information of space objects, and the model inference is illustrated to estimate the parameters of



shape and pose simultaneously. The experiments verify the effectiveness of our method and show the advantage of the approach proposed over the methods which rely on the known geometry of the object. Our research is valuable in exploration to understand the 3D pose and geometry of unknown objects from a single image.

## Data Availability

The underlying data supporting the results of our study is the public 3D models of space targets, which are from the Internet.

## Conflicts of Interest

The authors declare no conflict of interest.

## Acknowledgments

This work was supported by the National Key Laboratory Foundation (grant number 61404130316).

## References

- [1] T. Rybus, "Obstacle avoidance in space robotics: review of major challenges and proposed solutions," *Progress in Aerospace Sciences*, vol. 101, pp. 31–48, 2018.
- [2] F. Aghili, M. Kuryllo, G. Okouneva, and C. English, "Fault-tolerant pose estimation of space objects," in *2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 947–954, Montreal, ON, Canada, July 2010.
- [3] X.-H. Gao, B. Liang, Z.-H. L. Le Pan, and Y.-C. Zhang, "A monocular structured light vision method for pose determination of large non-cooperative satellites," *Journal of control, automation, and systems*, vol. 14, no. 6, pp. 1535–1549, 2016.
- [4] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "Pose estimation for spacecraft relative navigation using model-based algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 1, pp. 431–447, 2017.
- [5] B. Tamadazte, E. Marchand, S. Dembélé, and N. le Fort-Piat, "CAD model-based tracking and 3D visual-based control for MEMS microassembly," *The International Journal of Robotics Research*, vol. 29, no. 11, pp. 1416–1434, 2010.
- [6] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "Performance evaluation of 3d model-based techniques for autonomous pose initialization and tracking," *Communication Reports*, vol. 18, no. 1–2, pp. 55–63, 2013.
- [7] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "A model-based 3d template matching technique for pose acquisition of an uncooperative space object," *Sensors*, vol. 15, no. 3, pp. 6360–6382, 2015.
- [8] F. Terui, "Model based visual relative motion estimation and control of a spacecraft utilizing computer graphics," in *21st International Symposium on Space Flight dynamics, Toulouse, France*, pp. 1–15, Chofu-shi, Tokyo 182-8522, JAPAN, 2009.
- [9] L. Zhang, D. M. Wu, and Y. Ren, "Pose measurement for non-cooperative target based on visual information," *IEEE Access*, vol. 7, pp. 106179–106194, 2019.
- [10] X. Zhang, Z. Jiang, H. Zhang, and Q. Wei, "Vision-based pose estimation for textureless space objects by contour points matching," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 5, pp. 2342–2355, 2018.
- [11] S. Sharma, J. Ventura, and S. D'Amico, "Robust model-based monocular pose initialization for noncooperative spacecraft rendezvous," *Journal of Spacecraft and Rockets*, vol. 55, no. 6, pp. 1414–1429, 2018.
- [12] N. W. Oumer, S. Kriegel, H. Ali, and P. Reinartz, "Appearance learning for 3D pose detection of a satellite at close-range," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 125, pp. 1–15, 2017.
- [13] S. Sharma and S. D'Amico, "Comparative assessment of techniques for initial pose estimation using monocular vision," *Acta Astronautica*, vol. 123, pp. 435–445, 2016.
- [14] S. Sharma, B. Connor, and S. D'Amico, "Pose estimation for non-cooperative spacecraft rendezvous using convolutional neural networks," in *2018 IEEE Aerospace Conference*, pp. 1–12, Big Sky, MT, USA, Mar 2018.
- [15] V. Capuano, K. Kim, J. Hu, A. Harvard, and S. J. Chung, "Monocular-based pose determination of uncooperative known and unknown space objects," in *69th International Astronautical Congress (IAC)*, pp. 1–5, Bremen, Germany, 2018.
- [16] D. Ivanov, M. Ovchinnikov, and M. Sakovich, "Relative pose and inertia determination of unknown satellite using monocular vision," *International Journal of Aerospace Engineering*, vol. 2018, Article ID 9731512, 9 pages, 2018.
- [17] G. Pavlakos, X. Zhou, A. Chan, K. G. Derpanis, and K. Daniilidis, "6-DoF object pose from semantic keypoints," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, Singapore, May 2017.
- [18] X. Zhou, A. Karpur, L. Luo, and Q. Huang, "StarMap for category-agnostic keypoint and viewpoint estimation," in *Computer Vision – ECCV 2018*, vol. 11205 of Lecture Notes in Computer Science, pp. 328–345, Springer, Cham, 2018.
- [19] X. Zhou, M. Zhu, S. Leonardos, and K. Daniilidis, "Sparse representation for 3D shape estimation: a convex relaxation approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, pp. 1648–1661, 2017.
- [20] Y. Konishi, Y. Hanzawa, M. Kawade, and M. Hashimoto, "Fast 6D pose estimation from a monocular image using hierarchical pose trees," in *Computer Vision – ECCV 2016*, vol. 9905 of Lecture Notes in Computer Science, pp. 398–413, Springer, Cham, 2016.
- [21] N. G. Cho, A. Yuille, and S. W. Lee, "A novel linelet-based representation for line segment detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 5, pp. 1195–1208, 2018.
- [22] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks IEEE*, Perth, WA, Australia, Australia, Nov.-1 Dec. 1995.
- [23] J. J. Lim, H. Pirsiavash, and A. Torralba, "Parsing IKEA objects: fine pose estimation," in *2013 IEEE International Conference on Computer Vision*, pp. 2992–2999, Sydney, NSW, Australia, Dec 2013.
- [24] J. Xiao, B. Russell, and A. Torralba, "Localizing 3D cuboids in single-view images," in *Advances in neural information processing systems*, pp. 746–754, Neural Information Processing Systems Foundation, Inc., 2012.