

# 基于视觉的非合作空间目标三维姿态估计方法

任笑圆<sup>①</sup> 蒋李兵<sup>①</sup> 钟卫军<sup>②</sup> 王 壮<sup>\*①</sup>

<sup>①</sup>(国防科技大学ATR重点实验室 长沙 410073)

<sup>②</sup>(宇航动力学国家重点实验室 西安 724403)

**摘要:** 基于视觉的非合作空间目标三维姿态估计, 关键在于建立观测图像与目标模型的特征关联。当前方法往往通过采用复杂的多维特征、产生候选关联结果的方式确保特征关联的准确性, 难以兼顾算法效率。为解决以上问题, 该文提出一种结合深度学习技术的姿态估计方法, 首先通过深度神经网络得到姿态初值, 然后基于姿态初值建立图像和目标模型之间的特征关联, 进而求解目标姿态。所提方法中, 深度神经网络提供了稳定的姿态初值, 缩小了特征关联的候选空间; 在姿态初值的支撑下采取了更为高效的特征提取与匹配方法。仿真实验表明, 该文方法相比于现有方法更好地兼顾了算法准确率和效率。

**关键词:** 姿态估计; 空间目标; 深度学习; 特征匹配

中图分类号: TN911.73; TP391

文献标识码: A

文章编号: 1009-5896(2021)12-3476-10

DOI: 10.11999/JEIT200440

## A Vision-based Method for 3D Pose Estimation of Non-cooperative Space Target

REN Xiaoyuan<sup>①</sup> JIANG Libing<sup>①</sup> ZHONG Weijun<sup>②</sup> WANG Zhuang<sup>①</sup>

<sup>①</sup>(Key Laboratory on ATR, National University of Defense Technology, Changsha 410073, China)

<sup>②</sup>(State Key Laboratory of Astronautic Dynamics, Xi'an 724403, China)

**Abstract:** Establishing correspondence between the target model and the input image is an important step for the pose estimation of non-cooperative space target. Current methods always rely on complex image features and generation of candidate, which can be costly and time consuming. To solve the problems above, this paper proposes a pose estimation method that first conducts initial estimation based on deep neural network and then conducts accurate estimation through correspondence between the known target model and the input image is proposed. The deep neural network provides the stable initial value which reduces the candidates of correspondence between the target model and image. In addition, a more efficient feature extraction and matching method is adopted in this paper instead of complex multi-dimensional features. The simulation results show that the method proposed performs well both in efficiency and accuracy.

**Key words:** Pose estimation; Space target; Deep learning; Feature matching

### 1 引言

非合作空间目标的3维姿态估计是航天器在轨维护、太空碎片清除、目标跟踪监测等诸多航天应用的关键技术。与合作空间目标不同, 非合作空间目标无法直接提供姿态信息, 需要通过外部观测解算姿态。光学图像是空间目标监测的重要数据<sup>[1]</sup>, 本文研究基于单幅光学图像估计结构已知空间目标

的3维姿态。

当前, 基于单幅图像的空间目标姿态估计可大致分为基于模板和基于模型两种思路。基于模板的方法需要预先构建反映目标在不同姿态下表征的模板库, 通过输入图像与模板库的特征匹配确定目标姿态<sup>[2]</sup>。这类方法的精度受限于模板库中的模板数量, 而模板数量的增加会显著提高模板库构建成本、降低算法速度。基于模型的方法需要已知目标的3维模型, 通过目标3维模型和输入图像之间的特征关联计算目标的3维姿态<sup>[3-6]</sup>。相比于前一类方法, 此类方法不依赖大量的目标模板, 而是通过严格的特征关联方程求解得到目标准确的姿态值, 可取得较高的估计精度。此类方法的关键在于准确建

收稿日期: 2020-06-01; 改回日期: 2021-03-28; 网络出版: 2021-04-29

\*通信作者: 王壮 zhuang\_wang@sina.com

基金项目: 国防科技重点实验室基金(6142503180202)

Foundation Item: The Key Laboratory Foundation of National Defense Technology (6142503180202)

立输入图像和目标3维模型的特征关联。由于图像是3维空间的降维表达，因此在建立特征关联的过程中往往会产生较大的候选空间，当前方法通常需要筛选候选结果或者进行模型参数迭代来建立特征关联。Sharma等人<sup>[3]</sup>采用边缘特征和直线特征的组合得到3维模型和图像特征关联的候选结果，再从候选结果中筛选得到特征关联结果；Zhang等人<sup>[4]</sup>首先通过Hu矩特征匹配得到候选模型投影图像，然后通过观测图像与候选模型投影图像的轮廓特征匹配得到最终的特征关联。上述方法在特征关联过程中存在较大的候选空间，易造成特征的误匹配，降低算法效率和准确率。另外，为了保证特征关联的准确性，当前方法往往采取设计较为复杂的特征。例如，Zhang等人<sup>[4]</sup>采用ORB<sup>[7]</sup>算子进行轮廓上的点匹配，Rubio等人<sup>[5]</sup>所用到的SIFT<sup>[8]</sup>算子则需要计算复杂的128维特征，这些复杂的高维特征难以很好地兼顾算法效率。

随着深度学习技术的发展，Sharma等人<sup>[9]</sup>和李想<sup>[10]</sup>将卷积神经网络(Convolutional Neural Network, CNN)应用于空间目标的姿态估计，把姿态空间划分为不同的子空间，由此将姿态估计问题转换为基于深度神经网络的分类问题；也有学者基于神经网络直接建立2D-3D关联，进而求解姿态<sup>[11,12]</sup>。相比于基于传统人工特征的方法，神经网络可以从大量训练的数据中挖掘隐含特征，同时其样本数量会随着方法精度要求的提高而大幅增加。然而，对于空间目标来说，具有3维姿态标注的公开数据集十分有限，数据的缺乏限制了网络精度的提升。虽然提升CNN估计精度的训练成本较高，但基于CNN估计姿态的鲁棒性强，结果较为稳定，与上文提到的基于模型与图像的特征关联的姿态估计方法具有很好的互补性，结合以上两种方法的优势具有提升姿态估计准确率和效率的潜力。

针对现有方法在特征关联中存在的候选空间大、特征设计复杂的问题，本文提出一种结合深度学习和特征关联的姿态估计方法，首先基于CNN估计目标3维姿态，然后将其结果作为姿态初值建立图像与目标3维模型的特征关联，进而求解目标姿态的准确值。本文方法中CNN为后续的特征关联过程提供了稳定的初始值，相比于现有方法极大缩小了候选空间，降低了误匹配的可能性；在特征关联阶段，本文提出一种基于直线匹配的特征关联方法，相比于当前方法采用的复杂的高维特征拥有更高效的算法效率。

## 2 姿态估计方法

本文采用的参考坐标系定义与常用的光学摄像机成像模型一致，如图1(a)所示。目标3维姿态即为目标坐标系相对于相机坐标系的旋转角，具体为俯仰角 $\varphi$ 、翻滚角 $\gamma$ 和偏航角 $\theta$ ，定义如图1(b)所示。其中俯仰角 $\varphi$ 为 $X_T$ 轴与平面 $X_C O_C Y_C$ 的夹角，翻滚角 $\gamma$ 为通过 $X_T$ 轴的平面 $X_C O_C Y_C$ 的铅垂面与 $Z_T$ 轴的夹角，偏航角 $\theta$ 为 $X_T$ 轴在平面 $X_C O_C Y_C$ 上的投影和 $X_C$ 轴的夹角。

本文在已知目标模型的前提下根据单幅光学图像估计目标姿态，方法流程见图2，主要包括3个步骤：(1)采用CNN估计目标姿态，结果作为姿态初值；(2)将目标模型在姿态初值下的投影与输入图像进行特征匹配，建立目标模型与输入图像的特征关联，将目标姿态估计转换为PnP(Perspective-n-Point)问题，即给定3D空间和2D图像的对应点对求解位姿；(3)根据特征关联结果采取EPnP算法<sup>[13]</sup>求解目标姿态。考虑到太空特殊观测环境给特征提取与关联带来的干扰，在此步骤中评估姿态求解的偏差，若偏差过大则舍弃估计结果，采用基于CNN得到的姿态初值作为最终结果。下面详细介绍以上3个步骤。

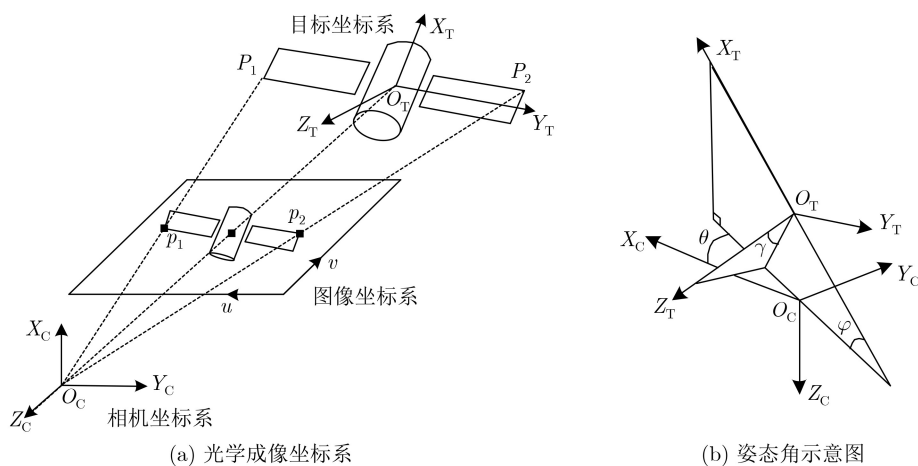


图1 坐标系及姿态角定义

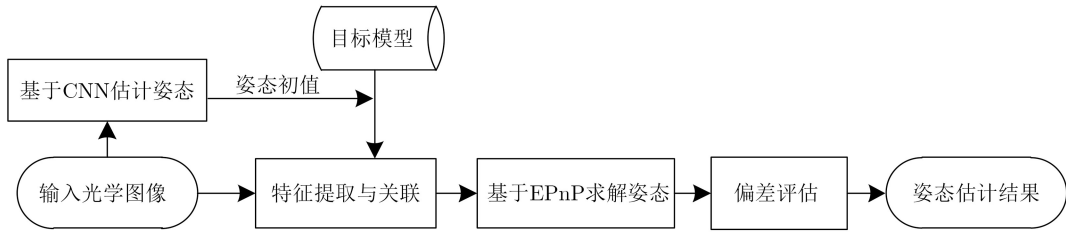


图2 方法流程图

2.1 基于CNN的姿态初值确定

本文通过CNN从图像中估计目标姿态，作为后续过程的姿态初值。目前PASCAL3D+<sup>[14]</sup>等公开3维数据集仅提供常见地物目标的图像样本，而具有3维信息标注的空间目标数据集十分缺乏。针对样本缺乏的问题，本文采用渲染目标CAD模型的方式生成训练数据，具体来说，使用Blender开源软件<sup>[15]</sup>得到不同姿态下的空间目标仿真图像作为网络的训练样本。为确保姿态空间划分的均匀性，本文采用球面等积网格<sup>[16]</sup>对姿态空间进行划分，在俯仰、翻滚和偏航3个维度上共划分4800个子空间。

本文在VGG<sup>[17]</sup>网络的基础上构建网络结构，如图3所示。其中卷积层与VGG网络的卷积层结构相同，在卷积层之后设置了3个全连接层。网络输入是224×224×3像素的RGB图像数据，输出是1×1×X的向量，代表姿态角的标签，对应于姿态

空间的特定子空间，X为子空间的数量。在网络中，所有卷积层的卷积核尺寸均为1×1，步长均为1；所有池化层的池化窗口尺寸均为2×2，步长均为2。

设 $Y_i$ 代表姿态的真实标签， $y_i$ 为对应的预测值，网络的损失函数定义为

$$\text{Loss}(Y, y) = - \sum_{i=1}^X Y_i \ln(y_i) \quad (1)$$

根据参数调优的结果，训练过程中的Batch Size设置为8，学习率设置为0.0001。

2.2 目标模型与输入图像的特征关联

空间目标往往由具有典型几何形状的部件组成，比如长方形太阳能帆板、立方体主体等，因此在观测图像中蕴含大量的直线特征。在CNN提供的初始姿态下，本文提取图像中的直线与目标模型投影进行匹配，进而根据直线端点得到匹配点对，

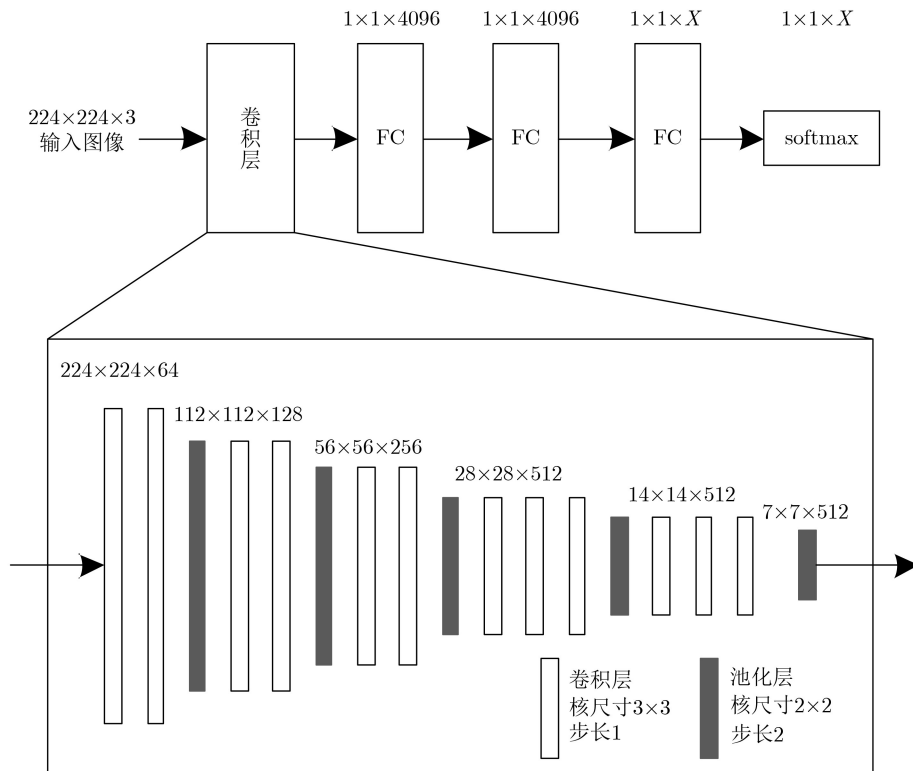


图3 姿态估计深度神经网络结构

完成目标模型和图像的特征关联。其中，目标模型为线框模型，其端点为典型几何形状的顶点，由目标CAD模型得到。特征匹配关联的具体流程如图4所示。

接下来介绍直线提取、直线匹配和关联点对获取方法。尽管空间目标图像中存在大量直线特征，然而图像中的明暗变换和目标零散部件给直线提取的完整性和准确性带来挑战。本文采用Cho等人<sup>[18]</sup>提出的直线提取方法，首先生成目标图像的梯度图，然后基于梯度图提取候选线段，进而将候选线段依据距离、斜率一致性、梯度变化程度等规则进行聚合，可有效连接由于零散部件和明暗变化导致断裂的直线。

令提取到的直线组表示为 $\mathbf{L} = \{L_i\}_N$ ，下一步将其与目标模型投影进行直线匹配。将目标线框模型在初始姿态下投影图像中的直线表示为 $\mathbf{L}^W = \{L_i^W\}_M$ ，目标模型投影到2维图像的过程与图1所示一致，可表示为

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2)$$

其中， $(x, y, z)$ 是目标在目标坐标系中的3维坐标， $\mathbf{R}$ 和 $\mathbf{t}$ 是目标坐标系相对于相机坐标系的旋转矩阵和平移量， $\mathbf{R}$ 为 $3 \times 3$ 的正交矩阵， $\mathbf{t}$ 为3维平移向量， $\mathbf{t} = (t_x, t_y, t_z)^T$ ，描述在3个坐标轴上的位移； $s$ 是缩放系数； $\mathbf{K}$ 为 $3 \times 4$ 的矩阵，表征相机的内参数， $\mathbf{0} = (0, 0, 0)$ 。其中， $\mathbf{R}$ 与3个姿态角的关系为

$$\mathbf{R} = \begin{pmatrix} \cos(\gamma) & 0 & \sin(\gamma) \\ 0 & 1 & 0 \\ -\sin(\gamma) & 0 & \cos(\gamma) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

接下来对直线组 $\mathbf{L}$ 和 $\mathbf{L}^T$ 进行匹配得到关联点对。与文献<sup>[19]</sup>类似，本文通过评估直线编组的相似程度实现直线匹配。令同一直线组中的两条直线

组成一直线对，设 $(L_a, L_b)$ 和 $(L_c^W, L_d^W)$ 分别为直线组 $\mathbf{L}$ 和 $\mathbf{L}^W$ 中的直线对，其不相相似度为

$$d(\{L_a, L_b\}, \{L_c^W, L_d^W\}) = \sqrt{\sum_k (f_k(L_a, L_b) - f_k(L_c^W, L_d^W))^2} \quad (4)$$

其中， $f_k(\cdot)$ 为直线对的关系函数。设 $(\mathbf{p}_1^a, \mathbf{p}_2^a)$ 和 $(\mathbf{p}_1^b, \mathbf{p}_2^b)$ 分别为直线 $L_a$ 和 $L_b$ 的两端点的坐标， $\mathbf{p}_{\text{mid}}^a$ 和 $\mathbf{p}_{\text{mid}}^b$ 分别为 $L_a$ 和 $L_b$ 的中点坐标，则直线 $L_a$ 和 $L_b$ 的关系函数 $f_k(L_a, L_b) = r_k(\mathbf{p}_1^a, \mathbf{p}_2^a, \mathbf{p}_1^b, \mathbf{p}_2^b)$ ，定义为

$$\left. \begin{aligned} r_1(\mathbf{p}_1^a, \mathbf{p}_2^a, \mathbf{p}_1^b, \mathbf{p}_2^b) &= \frac{|\mathbf{p}_1^a - \mathbf{p}_2^a|}{|\mathbf{p}_1^a - \mathbf{p}_2^a| + |\mathbf{p}_1^b - \mathbf{p}_2^b|} \\ r_2(\mathbf{p}_1^a, \mathbf{p}_2^a, \mathbf{p}_1^b, \mathbf{p}_2^b) &= \frac{|\mathbf{p}_{\text{mid}}^a - \mathbf{p}_{\text{mid}}^b|}{|\mathbf{p}_1^a - \mathbf{p}_2^a| + |\mathbf{p}_1^b - \mathbf{p}_2^b|} \\ r_3(\mathbf{p}_1^a, \mathbf{p}_2^a, \mathbf{p}_1^b, \mathbf{p}_2^b) &= \text{acos} \left( \text{abs} \left( \frac{(\mathbf{p}_1^a - \mathbf{p}_2^a) \cdot (\mathbf{p}_1^b - \mathbf{p}_2^b)}{|\mathbf{p}_1^a - \mathbf{p}_2^a| \cdot |\mathbf{p}_1^b - \mathbf{p}_2^b|} \right) \right) \end{aligned} \right\} \quad (5)$$

根据式(4)可得两直线组的不相似度集合 $\mathbf{D} = \{d_i(\{L_a, L_b\}, \{L_c^W, L_d^W\})\}$ ，其中 $a, b \in [1, N]$ ， $c, d \in [1, M]$ ；且 $a \neq b$ ， $c \neq d$ 。对集合中的元素由小到大进行排序，可得到有序集合 $\mathbf{D}_{\text{sort}} = \{d_i(\{L_a, L_b\}, \{L_c^W, L_d^W\}) | d_i \leq d_{i+1}\}$ ，其中 $i \in [1, N_D)$ ， $N_D$ 为集合 $\mathbf{D}_{\text{sort}}$ 中的元素数量。

得到直线对的集合后，根据 $\mathbf{D}_{\text{sort}}$ 确定直线组 $\mathbf{L}$ 和 $\mathbf{L}^W$ 之间的直线匹配。在 $\mathbf{D}_{\text{sort}}$ 中存在重复的匹配直线对，这里认为重复次数越多其权重越大。根据重复数量确定匹配直线对的权重，进而对 $\mathbf{D}_{\text{sort}}$ 中相互矛盾的匹配结果进行剔除，从而得到直线的匹配集合 $(\{L_1, L_2, \dots, L_C\}, \{L_1^W, L_2^W, \dots, L_C^W\})$ ，其中 $C$ 为匹配直线对的数量。将匹配集合中的直线端点作为匹配点，则可得到输入图像和目标模型投影图像之间的匹配点对记为 $(\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{2C}\}, \{\mathbf{p}_1^W, \mathbf{p}_2^W, \dots, \mathbf{p}_{2C}^W\})$ ，其中 $\mathbf{p}_i$ 和 $\mathbf{p}_i^W$ 为直线组 $\mathbf{L}$ 和 $\mathbf{L}^W$ 中的直线端点。得到匹配点对后需要进行筛选以移除异常值，这里通过引入匹配点分布的几何约束进行异常值去除，具体来说，首先在匹配点对集合中计算每对

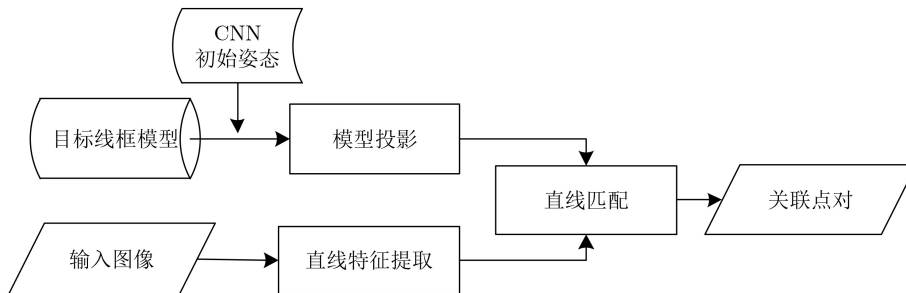


图4 特征提取与匹配关联示意图

匹配点的置信系数, 定义匹配点对 $\mathbf{p}_k$ 和 $\mathbf{p}_k^W$ 的置信系数

$$\phi(\mathbf{p}_k, \mathbf{p}_k^W) = \exp\left(-\sum_{i=1}^{2C} \sum_{j=1}^3 r_j(\mathbf{p}_k, \mathbf{p}_i, \mathbf{p}_k^W, \mathbf{p}_i^W)\right), \quad i \neq k \quad (6)$$

由此可得到匹配点置信系数集合 $\{\phi(\mathbf{p}_1, \mathbf{p}_1^W), \phi(\mathbf{p}_2, \mathbf{p}_2^W), \dots, \phi(\mathbf{p}_{2C}, \mathbf{p}_{2C}^W)\}$ , 选取其中大于阈值 $th$ 的置信所对应的点对作为最终的匹配点对。本文中阈值 $th$ 取 $0.1S$ ,  $S$ 为图像中目标最小绑定盒的最大边长。记最终得到的匹配点对为 $\{(\mathbf{p}_1, \mathbf{p}_1^W), (\mathbf{p}_2, \mathbf{p}_2^W), \dots, (\mathbf{p}_n, \mathbf{p}_n^W)\}$ 。设 $\mathbf{P}_i^W$ 为 $\mathbf{p}_i^W$ 对应的目标3维模型的点, 则可得图像与目标3维模型的匹配点对 $\{(\mathbf{p}_1, \mathbf{P}_1^W), (\mathbf{p}_2, \mathbf{P}_2^W), \dots, (\mathbf{p}_n, \mathbf{P}_n^W)\}$ 。

### 2.3 基于EPnP算法的姿态估计

2维图像和目标坐标系3维空间中的匹配点对记为 $\{(\mathbf{p}_1, \mathbf{P}_1^W), (\mathbf{p}_2, \mathbf{P}_2^W), \dots, (\mathbf{p}_n, \mathbf{P}_n^W)\}$ , 根据匹配点对求解目标3维姿态是一个PnP问题, 这里采用EPnP算法<sup>[13]</sup>解决。EPnP算法是目前解决PnP问题的常用方法, 相比于其他方法运算效率高, 有助于提升本文方法姿态估计的算法效率。通过EPnP算法可得到每个2维点 $\mathbf{p}_i$ 在相机坐标系中对应的3维点 $\mathbf{P}_i$ 的坐标, 令

$$\mathbf{A} = \begin{bmatrix} \mathbf{P}_1^W - \mathbf{P}_{center}^W \\ \dots \\ \mathbf{P}_n^W - \mathbf{P}_{center}^W \end{bmatrix} \quad (7)$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{P}_1 - \mathbf{P}_{center} \\ \dots \\ \mathbf{P}_n - \mathbf{P}_{center} \end{bmatrix} \quad (8)$$

其中,  $\mathbf{P}_{center}^W$ 和 $\mathbf{P}_{center}$ 的定义为

$$\mathbf{P}_{center}^W = \frac{1}{n} \sum_{i=1}^n \mathbf{P}_i^W \quad (9)$$

$$\mathbf{P}_{center} = \frac{1}{n} \sum_{i=1}^n \mathbf{P}_i \quad (10)$$

下面基于SVD分解求解姿态, 定义 $\mathbf{B}^T \mathbf{A}$ 为矩阵 $\mathbf{H}$ , 计算 $\mathbf{H}$ 的SVD分解

$$\mathbf{H} = \mathbf{U} \sum \mathbf{V}^T \quad (11)$$

则目标坐标系相对于相机坐标系的旋转矩阵为

$$\mathbf{R} = \mathbf{U} \mathbf{V}^T \quad (12)$$

平移量为

$$\mathbf{t} = \mathbf{P}_{center}^W - \mathbf{R} \cdot \mathbf{P}_{center} \quad (13)$$

其中, 旋转矩阵 $\mathbf{R}$ 表示目标坐标系相对于相机坐标

系的旋转角度, 在本文定义里即为目标姿态。将3维点 $\mathbf{P}_i^W$ 和 $\mathbf{R}$ 、 $\mathbf{t}$ 代入式(2)可得到其对应的投影点坐标 $\mathbf{p}_i^W$ , 令其与2维点 $\mathbf{p}_i$ 之间的偏差记为 $E$ , 定义为

$$E = \sum_{i=1}^n |\mathbf{p}_i^W - \mathbf{p}_i| \quad (14)$$

空间目标图像易产生高度降质, 可能对直线提取以及关联点的提取造成较大影响。为避免论文所提方法的特征关联阶段给方法准确性带来负面效果, 在此对偏差 $E$ 做评估, 若 $E$ 大于一定阈值, 则舍弃基于特征关联得到的姿态估计结果, 采用深度学习得到的整体姿态作为最终姿态结果。

### 3 仿真实验及分析

由于非合作空间目标的真实姿态无法获知, 且实测数据的获取较为困难, 因此本文采取仿真图像验证方法的有效性。在实验验证中, 本文选取了3个具有代表性结构的空目标, 分别为嫦娥一号卫星、天链一号卫星和风云三号卫星, 如图5所示。

在本文方法的网络训练阶段, 训练样本为目标CAD模型通过Blender开源软件渲染得到的图像。具体来说, 对每个目标在姿态区间 $\{\varphi \in (-180^\circ, 180^\circ), \gamma \in (-180^\circ, 180^\circ], \theta \in (-180^\circ, 180^\circ)\}$ 内对各个角度每隔 $9^\circ$ 进行采样, 产生 $8.4 \times 10^4$ 幅图像。

在特征关联阶段, 需要已知目标的3维线框模型。3维线框模型由目标CAD模型标注得到, 本文实验中的3个目标的线框模型如图6所示。

实验验证分为3个部分。3.1节给出所提方法进行空间目标姿态估计的定性和定量结果, 并分析所提方法中深度学习提供的初值对结果的影响; 3.2节将本文所提方法与基于传统特征建立2D-3D关联的方法进行比较; 3.3节将本文所提方法与通过神经网络直接建立2D-3D关联的方法进行比较。

#### 3.1 姿态估计结果及分析

首先给出本文方法的定性结果。以嫦娥一号为例, 本文方法的中间过程和定性结果如图7所示。

在图7中, 输入图像首先通过CNN得到姿态初始化结果, 表示为图中方框的朝向; 然后在输入图像中提取直线特征, 并与目标线框模型在初始姿态下的投影进行直线匹配, 得到特征关联结果; 接着, 在特征关联结果中提取直线的端点与模型投影中相应的点作为匹配点对, 并基于EPnP算法求解目标姿态, 得到姿态估计准确值。在图中, 姿态估计准确值对应的目标模型投影叠加在图像中。

其他两类目标的实验定性结果如图8和图9所示。第1列为基于CNN的姿态初值估计结果, 方框

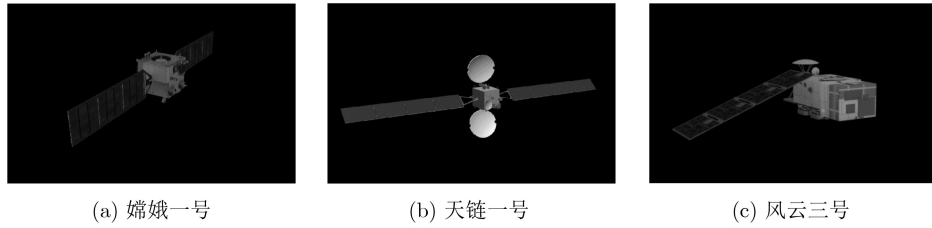


图5 实验验证目标示意

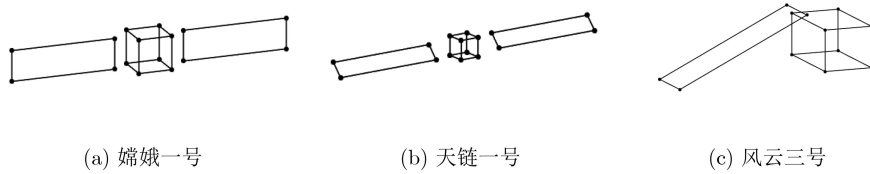


图6 实验目标线框模型

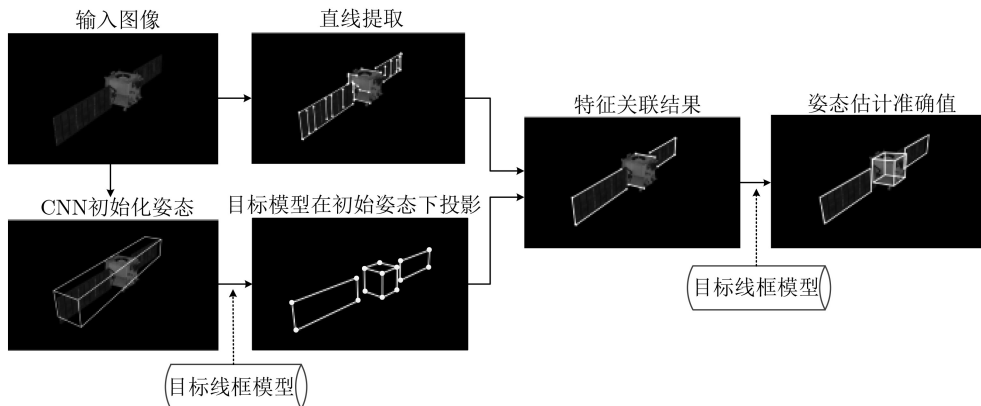


图7 本文方法中间过程和定性结果

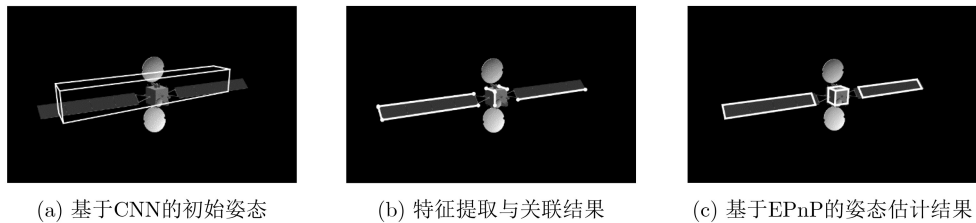


图8 “天链一号”姿态估计定性结果

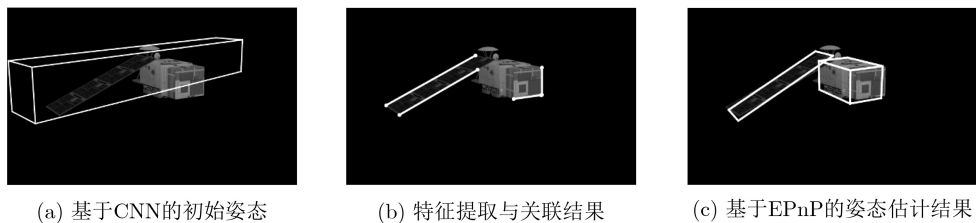


图9 “风云三号”姿态估计定性结果

的方向标示出估计得到的姿态；第2列为特征提取关联结果，图中标示出了与目标模型匹配的直线，其端点即为匹配点；第3列为采用EPnP算法得到的姿态估计准确值，图中标示出了在估计姿态下的目标模型投影。

下面给出定量结果。令姿态真值和估计值对应的旋转矩阵分别为 $\mathbf{R}_{true}$ 和 $\mathbf{R}$ ，则姿态估计误差表示为

$$\mathbf{R}_e = \mathbf{R} \cdot \mathbf{R}_{true}^T \quad (15)$$

令 $\mathbf{R}_e$ 对应的姿态误差角为 $\{\varphi_e, \gamma_e, \theta_e\}$ ，用以描

述姿态估计的准确性。令 $t_{true}$ 为平移量真值，方法估计值为 $t$ ，则平移量估计值的误差 $t_e$ 为

$$t_e = |t - t_{true}| \quad (16)$$

测试数据为3个测试目标的CAD模型渲染得到的仿真图像，每个目标各30幅。令 $\{\bar{\varphi}_e, \bar{\gamma}_e, \bar{\theta}_e\}$ 表示测试图像的平均姿态误差角， $\psi = (\bar{\varphi}_e + \bar{\gamma}_e + \bar{\theta}_e)/3$ 表示其3个角度的误差均值， $\bar{t}_e$ 为平移量误差均值。本文方法的定量结果如表1所示。

在本文方法中，CNN的姿态子空间数量会对估计精度和所需训练样本的数量产生影响，以嫦娥一号目标为例，图10给出不同姿态子空间下的估计精度和所需样本数量对比。

图10(a)是不同数量的姿态子空间下的网络所需的训练样本数，图10(b)是CNN姿态估计结果与在其基础上进行特征关联估计姿态的对比。从图10(a)可以看出，随着姿态子空间增多，所需样本数大幅增加；从图10(b)可以看出，CNN估计精度随姿态子空间增多而变高，但精度增幅相应地减小。由此可以看出对于基于CNN的姿态估计，随着精度需

表 1 本文方法姿态估计结果

	目标编号		
	(a)	(b)	(c)
$\{\bar{\varphi}_e, \bar{\gamma}_e, \bar{\theta}_e\}(\circ)$	{2.7, 2.9, 2.8}	{2.3, 2.2, 2.7}	{2.6, 2.5, 3}
$\psi(\circ)$	2.8	2.4	2.7
$\bar{t}_e(m)$	0.6	0.53	0.51

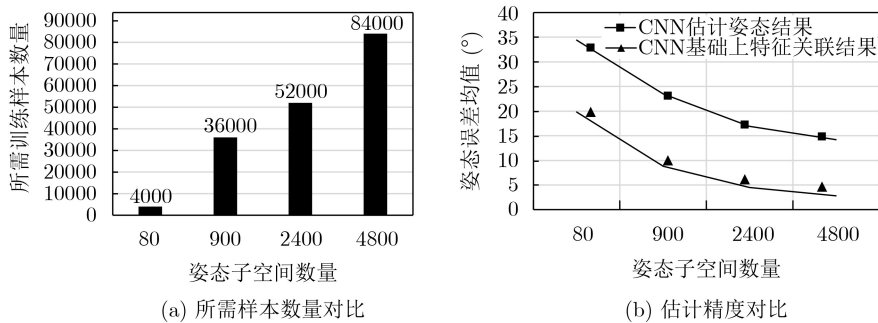


图 10 不同姿态子空间下的估计精度和所需样本数量对比

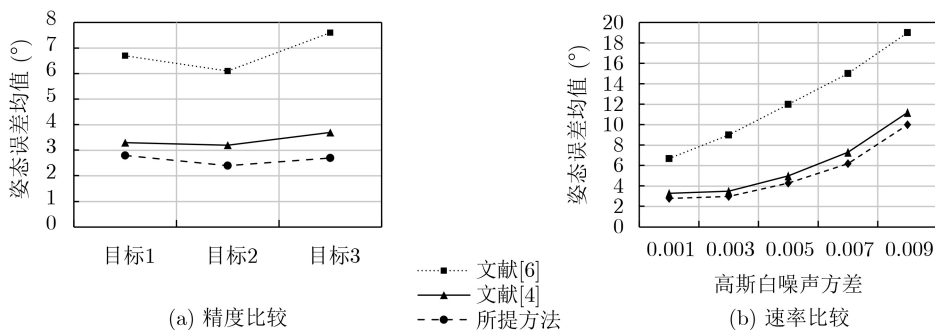


图 11 本文方法与基于传统特征方法的性能对比

求提升，训练成本将大幅提高。同时实验结果表明了在CNN基础上进行特征关联进而估计姿态的思路可显著提升估计精度。

### 3.2 与基于传统特征方法的对比

下面将本文方法与当前两种典型的基于传统特征实现2D-3D关联的方法进行对比。文献[6]在观测图像中提取空间目标的直线特征与太阳能帆板结构进行关联，从而求解姿态；文献[4]首先采用Hu特征矩生成候选空间，然后通过基于ORB算子的轮廓点匹配实现目标模型和图像的关联，进而求解目标姿态。将上述两种方法以及本文方法对3.1节的3个目标进行姿态估计，其算法精度和速率比较如图11所示。

可以看出文献[6]的方法有较高的算法效率，但其准确率较低；原因在于其所提方法仅通过帆板的直线特征建立特征关联，所提特征描述目标整体结构的能力有限。文献[4]使用了较为复杂的特征描述子进行特征关联，并构建候选空间以提供姿态初值，虽然方法准确率较高，但效率较低。本文方法在CNN提供初值的基础上通过特征关联求解姿态，相比于文献[6]，本文所提特征可较好地描述目标整体结构，在姿态估计准确率上有明显的优势；相比于文献[4]，本文方法通过CNN提供了稳定的姿态初值，有效缩小了特征提取的候选空间，并且本文提取特征相比于文献[4]采用的特征描述子复杂度低，从结果上看本文方法取得了与文献[4]相当的

准确率，并在算法效率上有较大优势。综合来看，本文方法较好地兼顾了准确率和算法效率。

由于太空环境的特殊性，观测图像往往存在降质。在空间观测环境中，噪声、低分辨、光照变化是导致图像降质的3种主要因素，下面分别测试本文方法在这3种典型的降质环境下的性能。将嫦娥一号作为测试对象，测试结果如图12所示。

在图像加噪测试中，高斯白噪声方差大于0.005后3种方法的误差显著增大。在图像低分辨测试中，目标在图像中所占像素数在10000以上时，所提方法可以较好地估计姿态，目标所占像素数为2000时，3种方法的误差显著增大。在光照变化测试中，光照相位角大于45°时对3种方法造成较大影响。从结果可以看出，图像退化对所有方法均产生了明显影响。本文方法和文献[4]都提取了目标整体特征进行姿态估计，文献[7]仅依赖目标局部特征，因此本文方法和文献[4]在图像退化的情况下取得了明显好于文献[7]的精度。

### 3.3 与基于深度学习直接建立2D-3D关联方法的对比

随着深度学习技术的发展，在学界有一部分方法通过网络直接获取目标的2D-3D匹配，进而解算目标姿态。这类方法依托公开的Pascal3D+<sup>[14]</sup>等3维数据集在常见地物目标(汽车、行人、家具等)的姿态估计中得到广泛应用。下面将本文方法

与一种基于深度学习直接建立2D-3D关联的方法进行对比<sup>[11]</sup>。文献[11]是通过网络直接建立2D-3D关联的具有代表性的方法，首先通过网络直接提取图像中目标的关键点建立2D-3D关联，然后根据关联结果求解得到姿态。这里通过渲染嫦娥一号的3Dmax模型得到数据集。本文方法和文献[11]方法对比如图13所示。

从结果可以看出，本文方法取得比文献[11]更高的姿态估计精度，特别是在图像退化的情况下。直接建立2D-3D特征关联的网络是从图像直接映射到目标的局部区域，然后通过数个局部区域的位置解算整体姿态。对于大部分的纹理丰富、成像清晰的目标，局部的关键点可以被准确提取。然而，空间目标多采用了弱纹理材质和重复纹理材质，加之空间目标图像可能存在高度降质，因此直接提取局部区域特征容易导致关键点提取的数量不足或提取错误，进而在整体姿态求解阶段陷入局部极小值或产生较大误差。本文采用一种“全局+局部”的思路，即首先通过深度学习整体估计目标的姿态，然后在整体姿态的基础上建立2D-3D关联。建立2D-3D关联后的姿态求解优化了整体姿态估计的结果，整体姿态估计结果又为建立2D-3D关联提供了稳定的初值，提高了关联的准确性。相比于直接映射到局部、建立2D-3D特征关联的方法，本文所提方法对于空间目标数据更加适用。

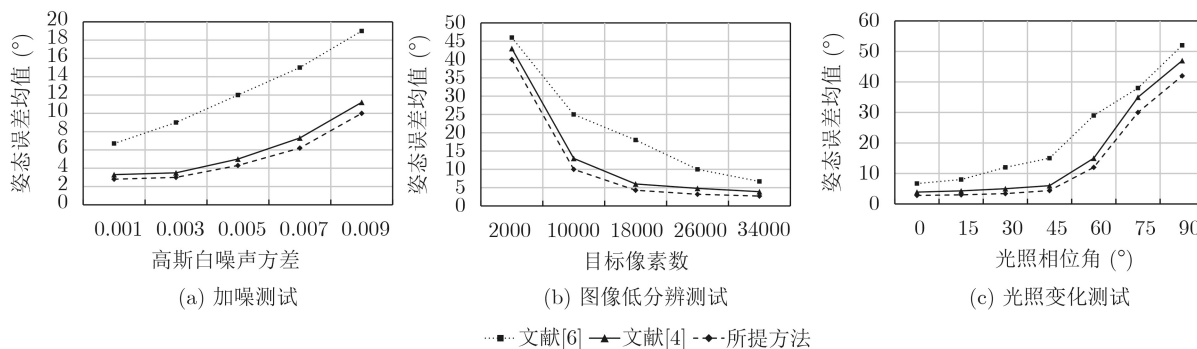


图 12 本文方法与基于传统特征方法在图像退化条件下的性能对比

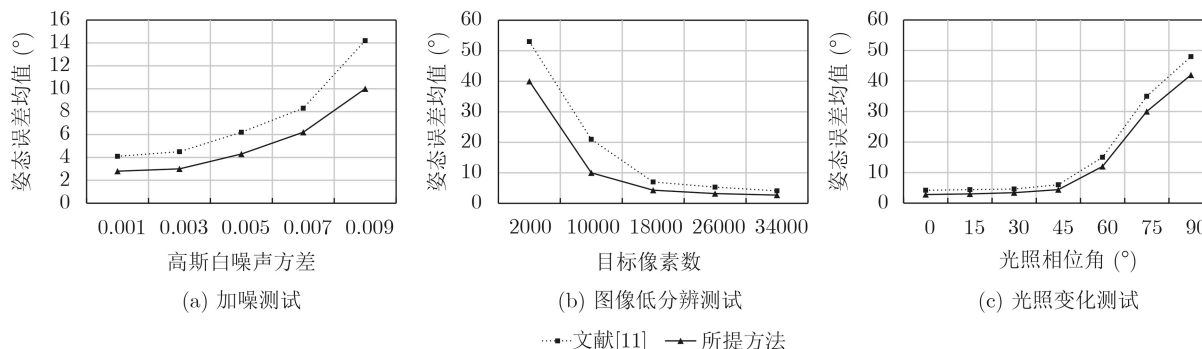


图 13 本文方法与基于深度学习方法在图像退化条件下的性能对比



## 4 结束语

本文提出一种基于视觉的非合作空间目标3维姿态估计方法,在已知目标3维模型的前提下估计单幅图像中空间目标的3维姿态。本文方法旨在解决现有方法在观测图像与目标3维模型特征关联中普遍存在的候选空间大、特征设计复杂的问题。首先基于CNN得到姿态初值,然后在姿态初值下建立观测图像与目标模型之间的特征关联,进而求解目标姿态的准确值。深度神经网络为特征关联过程提供了稳定的姿态初始值,有效缩小了候选空间。在特征关联阶段,本文提出了一种基于直线匹配的特征关联方法,相比于设计较为复杂的特征更为高效。相比于现有方法,实验证明所提方法相比于传统方法兼顾了准确率和效率,提升了非合作目标姿态估计的算法性能。

## 参考文献

- [1] 周叶剑,张磊,王虹现,等.多站ISAR空间目标姿态估计方法[J].电子与信息学报,2016,38(12):3182-3188. doi: 10.11999/JEIT160603.  
ZHOU Yejian, ZHANG Lei, WANG Hongxian, et al. Attitude estimation for space satellite targets with multistatic ISAR systems[J]. *Journal of Electronics & Information Technology*, 2016, 38(12): 3182-3188. doi: 10.11999/JEIT160603.
- [2] 王阳.基于轮廓线匹配的目标姿态求解研究与实现[D].[硕士论文],西安电子科技大学,2014.  
WANG Yang. Research and implementation of contour matching-based solution for target attitude[D]. [Master dissertation], Xidian University, 2014.
- [3] SHARMA S, VENTURA J, and D'AMICO S. Robust model-based monocular pose initialization for noncooperative spacecraft rendezvous[J]. *Journal of Spacecraft and Rockets*, 2018, 55(6): 1414-1429. doi: 10.2514/1.A34124.
- [4] ZHANG Xin, JIANG Zhiguo, ZHANG Haopeng, et al. Vision-based pose estimation for textureless space objects by contour points matching[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2018, 54(5): 2342-2355. doi: 10.1109/TAES.2018.2815879.
- [5] RUBIO A, VILLAMIZAR M, FERRAZ L, et al. Efficient monocular pose estimation for complex 3D models[C]. 2015 IEEE International Conference on Robotics and Automation, Seattle, USA, 2015: 1397-1402. doi: 10.1109/ICRA.2015.7139372.
- [6] 宋平,杨小冈,蔡光斌,等.单站光测图像中空间目标姿态估计[J].兵器装备工程学报,2020,41(9):165-170. doi: 10.11809/bqzbgcxb2020.09.031.  
SONG Ping, YANG Xiaogang, CAI Guangbin, et al. Attitude estimation for space object based on optical image[J]. *Journal of Ordnance Equipment Engineering*, 2020, 41(9): 165-170. doi: 10.11809/bqzbgcxb2020.09.031.
- [7] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[C]. 2011 International Conference on Computer Vision, Barcelona, Spain, 2011: 2564-2571. doi: 10.1109/ICCV.2011.6126544.
- [8] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110. doi: 10.1023/b:visi.0000029664.99615.94.
- [9] SHARMA S, BEIERLE C, and D'AMICO S. Pose estimation for non-cooperative spacecraft rendezvous using convolutional neural networks[C]. 2018 IEEE Aerospace Conference, Big Sky, USA, 2018. doi: 10.1109/AERO.2018.8396425.
- [10] 李想.基于深度学习的空间非合作目标姿态估计算法设计[D].[硕士论文],哈尔滨工业大学,2018.  
LI Xiang. Design of attitude estimation algorithm for space non-cooperative targets based on deep learning[D]. [Master dissertation], Harbin Institute of Technology, 2018.
- [11] PAVLAKOS G, ZHOU Xiaowei, CHAN A, et al. 6-DoF object pose from semantic keypoints[C]. 2017 IEEE International Conference on Robotics and Automation, Singapore, 2017. doi: 10.1109/ICRA.2017.7989233.
- [12] ZAKHAROV S, SHUGUROV I, and ILIC S. DPOD: 6D pose object detector and refiner[C]. 2019 IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019.
- [13] 张慧娟,熊芝,劳达宝,等.基于EPNP算法的单目视觉测量系统研究[J].红外与激光工程,2019,48(5):180-185.  
ZHANG Huijuan, XIONG Zhi, LAO Dabao, et al. Monocular vision measurement system based on EPNP algorithm[J]. *Infrared and Laser Engineering*, 2019, 48(5): 180-185.
- [14] XIANG Yu, MOTTAGHI R, and SAVARESE S. Beyond PASCAL: A benchmark for 3D object detection in the wild[C]. 2014 IEEE Winter Conference on Applications of Computer Vision, Steamboat Springs, USA, 2014. doi: 10.1109/WACV.2014.6836101.
- [15] 钱旭冉,朱明,蔡栩.基于Blender的天文数据三维可视化[J].天文研究与技术,2015,12(2):204-211. doi: 10.3969/j.issn.1672-7673.2015.02.010.  
QIAN Xuran, ZHU Ming, and CAI Xu. A study of 3D visualization of astronomical data based on the blender[J]. *Astronomical Research and Technology*, 2015, 12(2): 204-211. doi: 10.3969/j.issn.1672-7673.2015.02.010.
- [16] 贲进,童晓冲,张永生,等.球面等积网格系统生成算法与软件模型研究[J].测绘学报,2007,36(2):187-191. doi: 10.3321/j.issn:1001-

- 1595.2007.02.013.  
BEN Jin, TONG Xiaochong, ZHANG Yongsheng, *et al.* Research on generating algorithm and software model of discrete global grid systems[J]. *Acta Geodaetica et Cartographica Sinica*, 2007, 36(2): 187–191. doi: [10.3321/j.issn:1001-1595.2007.02.013](https://doi.org/10.3321/j.issn:1001-1595.2007.02.013).
- [17] SIMONYAN K and ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv: 1409.1556, 2014.
- [18] CHO N G, YUILLE A, LEE S W, *et al.* A novel linelet-based representation for line segment detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(5): 1195–1280. doi: [10.1109/TPAMI.2017.2703841](https://doi.org/10.1109/TPAMI.2017.2703841).
- [19] 胡海霞, 李钢. 几何特性二元关系的直线匹配[J]. *中国图象图形学报*, 2014, 19(9): 1338–1348. doi: [10.11834/jig.20140911](https://doi.org/10.11834/jig.20140911).
- HU Haixia and LI Gang. Line matching based on binary relations of geometric attributes[J]. *Journal of Image and Graphics*, 2014, 19(9): 1338–1348. doi: [10.11834/jig.20140911](https://doi.org/10.11834/jig.20140911).
- 任笑圆: 男, 1991年生, 博士生, 研究方向为目标检测识别、图像融合.
- 蒋李兵: 男, 1982年生, 讲师, 研究方向为SAR图像解译、图像融合.
- 钟卫军: 男, 1982年生, 工程师, 研究方向为空间目标探测与识别.
- 王 壮: 男, 1973年生, 教授, 研究方向为雷达信息处理、空间目标监视、目标识别.

责任编辑: 马秀强