

# 双模态信息融合的飞行目标位姿估计方法

李荣华<sup>1</sup>, 王 蒙<sup>1</sup>, 周 唯<sup>2</sup>, 付佳茹<sup>1</sup>

(1. 大连交通大学 机械工程学院, 辽宁 大连 116028;

2. 中国人民解放军 91550 部队, 辽宁 大连 116023)

**摘要:** 针对在飞行目标位姿估计中背景复杂、目标提取及位姿解算精度低、速度实时性差等问题, 提出了一种激光与图像信息融合的飞行目标位姿估计方法。首先, 建立彩色相机和激光雷达间的坐标变换模型实现两传感器像素级匹配, 对同一时刻的图像和点云进行融合处理; 其次, 采用 ViBe 算法与深度信息融合对图像中运动目标进行提取, 并根据图像目标的位置框选出对应的点云; 最后, 利用 PnP 算法进行特征点粗配准, 获取点云间初始旋转平移矩阵, 并采用迭代最近点算法进行精配准, 利用 IK-Tree 加速临近点搜索, 提高配准速度。使用仿真试验和半实物仿真试验对方法准确性和稳定性进行了验证, 结果显示: 二维图像目标检测算法正确率为 97%, 错误分类比为 0.0112%; 位姿估计算法与传统迭代最近点算法相比精度提高了 53.2%, 单次耗时从 261 ms 降低至 132 ms, 效率提升约 49.4%, 与其他算法相比也具有一定优势。为飞行目标落点精准预测和制导控制提供解决思路。

**关键词:** 双模态信息融合; 位姿估计; 飞行目标; 目标提取

**中图分类号:** TP391 **文献标志码:** A **DOI:** 10.3788/IRLA20220618

## 0 引 言

随着信息化、智能化技术的不断发展, 现代战争对武器装备打击的精准度要求越来越高。导弹飞行试验中实时解算导弹位姿、准确预报导弹落点有利于判断导弹是否击中目标, 并且可以及时发现导弹故障实施提前炸毁<sup>[1]</sup>。与传统的建立动力学、运动学数学模型推测导弹落点<sup>[2]</sup>不同, 如今通过对飞行过程中导弹位姿信息进行捕捉, 并根据抛物线近似、非线性滤波、深度学习等方法预测导弹落点<sup>[3-4]</sup>, 符合信息化、智能化的要求。因此, 实现对复杂背景下的飞行目标实现快速位姿估计, 成为了目前亟需解决和攻克的问题。

国内外学者针对复杂背景下飞行目标检测、位姿估计问题提出大量研究方法。例如, 乐英等<sup>[5]</sup>等提出了基于背景差分 and 帧间差分法对目标进行检测, 增强了背景差分法的检测能力, 使提取的运动目标相对完

整且噪声少。Barnic 等<sup>[6]</sup>提出 ViBe (Visual Background Extractor) 算法对运动目标进行检测, 利用图像初始帧来构建背景模型, 计算量小且实时性高, 但是当初始帧包含运动目标时, 容易将运动目标初始化到背景模型中产生伪前景。刘家军等<sup>[7]</sup>针对 ViBe 算法检测运动目标时产生伪前景的问题, 在 ViBe 算法的基础上使用熵值法判断颜色和空间位置在相似程度函数中的权重, 采用迭代法确定自适应阈值, 以增强前景分割精度, 但 ViBe 算法容易受光照影响, 在面临光照变化、动态干扰等情况时检测性能较差。胡涛<sup>[8]</sup>等提出 RGBD 相机与 ViBe 算法结合, 将深度信息作为背景建模的样本, 提高了 ViBe 算法的稳定性, 但该算法无法直接应用于彩色相机分辨率和雷达分辨率不同的情况。尹芳<sup>[9]</sup>采用 ICP 算法计算相邻关键帧之间的相对位姿信息, 但对于飞行速度快的目标, ICP 配准算法无法处理初始位置相差较大的点云并易陷入

收稿日期: 2022-08-10; 修订日期: 2022-09-15

基金项目: 国家自然科学基金青年科学基金 (41901415); 国防科技重点实验室基金项目 (2022-JCJQ-L8-015-0201); 辽宁省教育厅科学研究项目重点项目 (LJKZ0475); 大连市杰出青年科技人才项目 (2022RJ03)

作者简介: 李荣华, 男, 教授, 博士生导师, 博士, 主要从空间非合作目标构型重建与测量技术研究等方面的研究。

局部最优。田光宝等<sup>[10]</sup>采用 3 个不共线的特征点对目标进行位姿求解,但在特征点提取时容易产生较大误差影响位姿估计精度;刘耀华等<sup>[11]</sup>利用聚类算法和随机森林算法对点云进行分割,然后对分割后的实例采用 FLANN 算法进行特征匹配得到实例分割后特征点所对应的匹配点。最后通过 RANSAC 算法和 ICP 算法求解目标的精确位姿,但该算法不适用于背景复杂且目标物小的场景。Kechagias<sup>[12]</sup>提出了激光雷达和相机融合的方法估计目标位姿,利用点云与 RNN 网络结合进行目标相对姿态估计。但飞行目标视觉特征不明显,在训练过程中卷积神经网络容易被忽略,且目标在飞行过程中存在尺度变化幅度大的情况,也给训练带来了很大困难。尽管上述方法可以有效对目标进行检测和位姿估计,但在导弹落点精准预测和制导控制上存在无法在复杂背景下对飞行目标快速、准确提取以及位姿估计的问题。

三维测量中线阵雷达主要依赖线列推扫式,其优势在于灵活的视场配置和较高的点云密度,但受限于扫描速率,在对动态目标进行观测时,其测量结果与实际轮廓存在较大差异。而基于高灵敏度阵列的面阵激光雷达,可实现单脉冲瞬时成像,有效解决了运动扫描导致的时间同步问题和条带错位现象<sup>[13]</sup>。

因此,在保证实时性的前提下,文中提出基于面阵式激光雷达与图像信息融合的双模态飞行目标位姿估计方法。首先,采用 ViBe 算法与深度信息结合,实现图像目标快速提取并且可以较好的抑制噪点;其次,通过将整体点云和图像目标融合,实现从整体点云中提取出目标点云;最后,采用 PnP (Perspective-n-Point) 算法获取点云间初始旋转平移矩阵并采用基于 IK-D Tree (Incremental K-dimensional Tree) 的 ICP (Iterative Close Point) 算法解算精确位姿。文中方法与现有方法相比计算速度快、提取目标精度高,能够很好应对背景复杂的场景特殊性。

### 1 双模态融合目标信息提取算法

文中所提方法的总体框架包含两个部分:图像目标提取部分和点云信息处理部分,其中点云信息处理包含特征叠加和目标点云提取,其流程如图 1 所示。

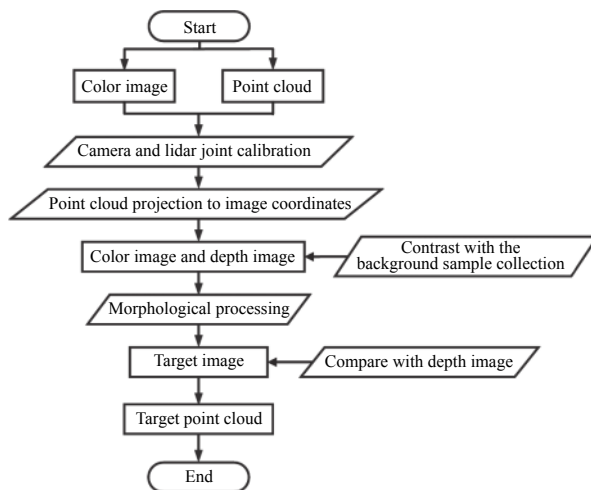


图 1 双模态融合目标信息提取算法流程图

Fig.1 Flow chart of bimodal fusion algorithm

#### 1.1 激光与可见光相机联合标定

研究中所用彩色相机的出厂基本参数不能满足高精度估计要求,且彩色相机在成像过程中会因镜头制造误差和组装误差等原因导致成像效果不理想。因此,采用张正友相机标定法对彩色相机内参进行重新标定。标定完成后采用 PnP 算法求解激光雷达至彩色相机的映射矩阵。具体方法为:用彩色相机和激光雷达对同一特征目标进行拍摄,采用手动标记的方式标记出特征点在彩色相机和点云中的位置。假设标记的特征点在激光雷达坐标系下的坐标为  $(X_L, Y_L, Z_L)$ , 在图像像素坐标系下的坐标为  $(u, v)$  可建立如下关系:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = Q \begin{bmatrix} R_{L-C} & T_{L-C} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_L \\ Y_L \\ Z_L \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_L \\ Y_L \\ Z_L \\ 1 \end{bmatrix} \quad (1)$$

式中:  $Q$  表示彩色相机内参;  $R_{L-C}$  表示激光雷达坐标系到彩色相机坐标系的旋转矩阵;  $T_{L-C}$  表示平移矢量;  $m_{11} \sim m_{34}$  表示待求矩阵的 12 个元素。

通过提取多组激光雷达与图像对应特征点,建立如下约束方程组,即可求解出矩阵的 12 个元素。由公式 (1) 展开可得到:

$$\begin{cases} u = m_{11}X_L + m_{12}Y_L + m_{13}Z_L + m_{14} \\ v = m_{21}X_L + m_{22}Y_L + m_{23}Z_L + m_{24} \\ 1 = m_{31}X_L + m_{32}Y_L + m_{33}Z_L + m_{34} \end{cases} \quad (2)$$

求解公式 (2) 至少需要 6 组特征点,为减小误差

采集多组特征点进行解算。应用最小二乘法求得映射矩阵初始解,并用 Levenberg-Marquarat 算法对其结果进行优化得到最终结果。

### 1.2 激光与可见光融合方法

求解出彩色相机内参和激光雷达至彩色相机的映射矩阵后,通过转换矩阵实现点云坐标到图像像素坐标的投影。为了实现点云坐标与图像像素坐标的相互转换,将各变换矩阵扩展为 4×4 的方阵。若将点云坐标表示为  $P(X_L, Y_L, Z_L)$ , 映射到图像像素坐标系的二维映射点云表示为  $Y(u, v)$ , 则点云坐标到图像像素坐标的映射计算方式如下:

$$M_0 = \begin{bmatrix} R_{L-C} & T_{L-C} \\ 0 & 1 \end{bmatrix} \quad (3)$$

$$Z_L \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} Q & 0 \\ 0 & 1 \end{bmatrix} \cdot M_0 \cdot \begin{bmatrix} X_L \\ Y_L \\ Z_L \\ 1 \end{bmatrix} \quad (4)$$

式中:  $M_0$  为扩充后的外参矩阵。

将点云数据转换至图像像素坐标系,即可完成点云数据与图像数据的特征叠加。点云与图像融合效果如图 2 所示。



图 2 点云与图像融合

Fig.2 Point cloud and image fusion

### 1.3 基于 ViBe 算法与深度信息融合的图像目标提取算法

针对在传统 ViBe 算法中存在的因光照强度改变而影响算法精度的问题,利用激光雷达不受光照影响的优势,采用 ViBe 算法与深度信息融合,将深度信息加入背景模型样本中作为运动目标检测的权值。

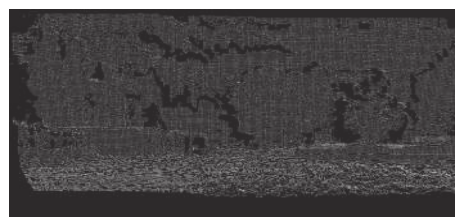
#### 1.3.1 稠密深度信息背景建模

文中采用将点云信息通过坐标转换投影至图像上的方式赋予每个像素点深度值。但受到当前传感器芯片技术水平的限制,直接采集到的三维点云数据通常分辨率较低,远不及可见光图像<sup>[14]</sup>,无法做到给

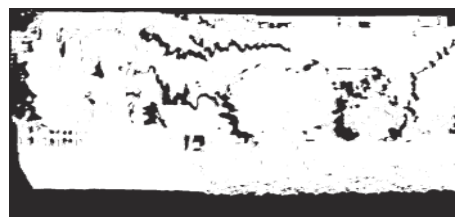
每个像素赋予深度信息。因此需要对点云进行稠密化。

由面阵式激光雷达扫描成像原理可知,点云在空间的分布情况为呈锥形方向均匀展开。因此通过视场角对相同物距的某一平面激光雷达可视区域的 4 个角点计算,即可得出雷达在此平面的最大投影面积<sup>[15]</sup>。假设雷达到测量面的距离为  $L$ , 彩色相机分辨率为:  $a_C \times b_C$ , 雷达分辨率为  $a_L \times b_L$ , 雷达水平视场角为  $H$ , 垂直视场角为  $G$ , 4 个角点的分别为  $P_1, P_2, P_3, P_4$ , 4 个角点的坐标为:  $P_1(-L \cdot \tan(\frac{H}{2}), L \cdot \tan(\frac{G}{2}), L)$ ,  $P_2(L \cdot \tan(\frac{H}{2}), L \cdot \tan(\frac{G}{2}), L)$ ,  $P_3(-L \cdot \tan(\frac{H}{2}), -L \cdot \tan(\frac{G}{2}), L)$ ,  $P_4(L \cdot \tan(\frac{H}{2}), -L \cdot \tan(\frac{G}{2}), L)$ 。假设  $P_1, P_2, P_3, P_4$  对应像素坐标分别为:  $Y_1(u_1, v_1), Y_2(u_2, v_2), Y_3(u_3, v_3), Y_4(u_4, v_4)$ , 则点云转化至图像坐标系后,投影最大面积为  $4 \cdot u_2 \cdot v_2$ , 深度信息在彩色图像上的占比率为  $V$  为  $\frac{4 \cdot u_2 \cdot v_2}{a_L \cdot b_L}$ 。

在实际应用中由于遮挡物(地面、天花板)等原因会导致映射至图像坐标系后,某一部分区域深度信息在彩色图像上的占比会大于  $V$ , 图像中一个像素点周围有  $(2r+1)^2 - 1$  个像素(其中  $r$  为最远处像素距中心像素的水平距离),因此将不小于  $V$  的最小  $(2r+1)^2 - 1$  作为深度信息填充的阈值,填充效果如图 3 所示。



(a) 原始深度图  
(a) Original depth map



(b) 稠密深度图  
(b) Dense depth map

图 3 点云稠密化

Fig.3 Point cloud densification

填充方法如下:

- (1) 建立零矩阵  $M(a_C \cdot b_C)$ , 将点云投影至彩色相机

坐标系后对坐标信息进行向下取整,根据坐标将深度值填充至  $M$  中。

(2) 寻找出  $M$  中的每个零点坐标,在周围距离为  $r$  的像素内寻找非零点,将所有非零点的均值作为该点的深度值。

在稠密深度信息构建完成后,使用三张背景图作为背景模型样本集合。在三张背景图的每一像素的 8 个相关邻域中随机选取 5 个像素值组成该坐标的像素背景模型;然后,对三张背景图对应的点云信息进行 1.3.1 节中所说的融合、稠密化处理,并选取对应坐标的深度值作为该坐标的深度背景模型。

### 1.3.2 基于最大类间方差法的自适应阈值选取

背景建模完成后,需要根据阈值区分前景和背景。文中采用参考文献 [6] 所提算法获取像素确定自适应阈值,采用最大类间方差法获取深度信息确定自适应阈值。

首先使用背景深度矩阵减去含运动目标的深度矩阵并向下取整,记为深度差值矩阵  $U$ ,并查找矩阵  $U$  中最大的深度差值,记为  $L$ 。然后计算矩阵  $U$  中深度值为  $i$  的点在矩阵中的分布概率  $P_i$ ,其中:  $P_i = \frac{U \text{中深度差值为} i \text{的点数}}{U \text{总点数}}$ , ( $i=1, 2, 3 \dots L$ )。则最大类间方差构建方法为:

$$P_1 = \sum_{i=0}^k P_i \quad (5)$$

$$m = \sum_{i=0}^k iP_i \quad (6)$$

$$mG = \sum_{i=0}^{L-1} iP_i \quad (7)$$

$$\sigma^2 = \frac{(mG \times P_1 - m)^2}{P_1 \times (1 - P_1)} \quad (8)$$

式中:  $P_i$  为矩阵  $G$  中每个深度值在矩阵中的分布概率, ( $i=1, 2, 3 \dots L$ );  $k$  为所求自适应阈值;  $m$  为灰度级  $k$  的累加均值;  $mG$  为图像全局均值。

将公式 (5)~(7) 代入公式 (8) 中,求解出令公式 (8) 最大化的差值  $k$  即为最大类间方差法的阈值。

### 1.3.3 图像目标提取

在实际应用中,由于扫描设备自身的限制或待测地形(物)的特点,获取的点云数据中往往会有大量的空洞<sup>[16]</sup>。为避免空洞点对整体算法的影响,当某一像

素点对应的深度背景模型深度值为 0 时仅使用像素值作为运动目标检测的权值。在图像目标提取完成时采用中值滤波去除图像中的离群噪点,提取效果如图 4 所示。



图 4 运动目标提取  
Fig.4 Moving object extraction

由于目标表面像素值与背景像素值接近、目标表面由于镜面反射成像后颜色信息发生改变等原因,导致提取出的目标有缺陷、不完整。因此使用图像形态学中的腐蚀和膨胀操作对目标缺陷部分进行填充。先采用膨胀操作对二值化物体进行扩充,再采用腐蚀操作对图像大小进行还原即可得到完整的运动目标,效果如图 5 所示。

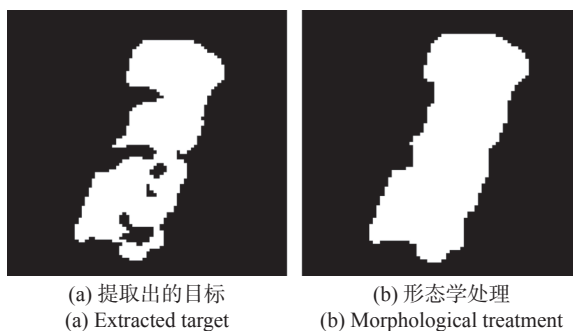


图 5 图像形态学处理

Fig.5 Image morphological processing

### 1.4 点云目标提取

对 1.3 节中求解到的目标图像使用“sobel”算子

提取目标边缘轮廓的像素坐标,并将点云坐标投影至图像像素坐标。通过比对目标区域和投影点云坐标去判断出目标投影点云的区域。

在点云投影至图像像素坐标系后,点云的相对位置并未改变,因此通过观察目标投影点云在整体投影点云的位置即可得到目标三维点云在原始点云中的位置,效果如图 6(a) 所示。在框选目标投影点云范围选择点云时在边缘部分容易提取到背景的点云信息,因此最后对提取出的点云进行去噪处理即可得到目标的三维点云,效果如图 6(b) 所示。

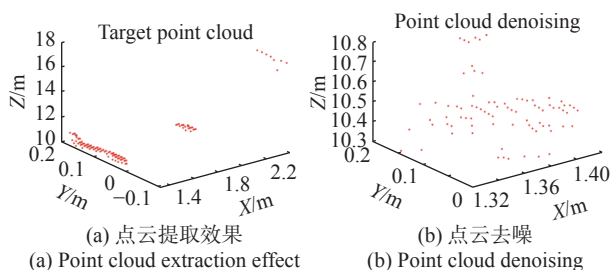


图 6 目标点云提取

Fig.6 Target point cloud extraction

## 2 双模态融合的改进 ICP 算法

使用 ICP 算法对飞行目标进行位姿估计时,存在以下两个问题:

(1) 由于提取到的点云目标与点云模型之间距离相差大。传统 ICP 算法在两个待配点云欧式距离较远时,会造成匹配错误,影响最终匹配效果;

(2) 传统 ICP 算法进行最近点查找时需计算欧氏距离最短的点。查找速度慢,配准效率低。

针对该问题,采用基于图像和点云特征点关系的 PnP 算法获取点云间初始旋转平移矩阵,完成点云目标与点云模型粗配准。采用 IK-D Tree 优化加速临近点搜索。

### 2.1 PnP 算法粗配准

对 1.3 节中得到的运动区域采用 LSD 直线检测算法检测直线段,然后采用霍夫变换对直线进行提取。在霍夫变换时存在一条直线段被提取成多条直线段的情况。因此,根据直线段到坐标原点的距离和直线段与坐标原点的夹角将直线段划分为 4 类,将划分在同一组的直线采用最小二乘法进行直线拟合,从

而得到所需直线,特征直线提取如图 7 所示。最后,通过计算直线间的交点得到 4 个目标特征点的图像像素坐标。

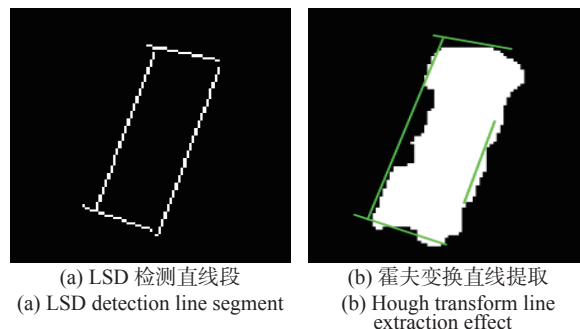


图 7 特征直线求取

Fig.7 Feature line extraction

根据目标真实尺寸采用相关软件进行目标三维点云模型创建。以模型几何中心作为坐标原点建立物体坐标系。采用手动标点的方式得到点云模型的 4 个特征点。待测目标物体模型及其坐标系定义如图 8 所示。已知 4 个特征点在图像像素坐标系和在物体坐标系下的坐标,利用 PnP 算法计算出目标在图像像素坐标系下的位姿信息,粗配准效果如图 9 所示。

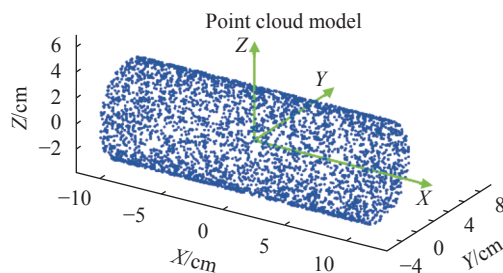


图 8 点云模型坐标系

Fig.8 Point cloud model coordinate system

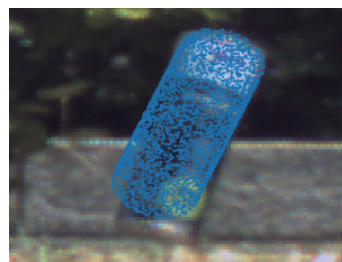


图 9 位姿求解结果

Fig.9 Pose solution results

假设求得的旋转矩阵和平移矩阵分别为:  $R_{cp}(3 \times 3)$  和  $T_{cp}(1 \times 3)$ , 目标原始点云在物体坐标系下的

坐标为  $P(X_M, Y_M, Z_M)$ , 目标在图像像素坐标系下的点云为  $P_F(X_F, Y_F, Z_F)$ 。对各变换矩阵全部扩展成  $4 \times 4$  的方阵, 则计算公式为:

$$M_1 = \begin{bmatrix} R_{cp} & T_{cp} \\ 0 & 1 \end{bmatrix} \quad (9)$$

$$\begin{bmatrix} X_F \\ Y_F \\ Z_F \\ 1 \end{bmatrix} = M_1 \cdot \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} \quad (10)$$

最后, 利用公式 (4) 将点云  $P_{Fi}$  从图像像素坐标系转换至激光雷达坐标系, 即完成点云粗配准。

## 2.2 优化 ICP 精配准

### 2.2.1 优化最近点搜索

采用 PnP 算法获得粗略六维位姿后, 模型点云与目标点云已大致重合, 但精度较低。为提高配准精度, 采用基于 IK-D Tree<sup>[17]</sup> 的 ICP 算法实现点云精配准。IK-D Tree 加速对近邻点搜索构建思路如下:

(1) 按照模型点云中点集的最大协方差构建传统 K-D Tree, 并用 K 近邻搜索 (k Nearest Neighbor, kNN) 查询点距离最近点;

(2) 对未匹配的子树采用 Lazy LabelInit 进行删除, 并使用 pullup 函数将未匹配的子树信息储存到节点  $T_i(i=1,2,3 \cdots n)$ 。其中包括 treesize (子树的尺寸); invalidnum (子树上被标记为“deleted”的节点数字), 以及 range (以当前节点为根节点的子树中, 所有节点在三维空间中的最小包围盒), 完成 IK-D Tree 构建。

(3) 检测 IK-D Tree 的平衡性。平衡准则由两个子准则组成:  $\alpha$ -平衡准则和  $\alpha$ -删除标准。假设 IK-D tree 的一个子树的根在  $T_n$ 。当子树满足以下条件:

$$S(T_n.leftson) < \alpha_{bal}(S(T_n) - 1) \quad (11)$$

$$\begin{aligned} S(T_n.rightson) < \alpha_{bal}(S(T_n) - 1) \\ I(T_n) < \alpha_{del}S(T_n) \end{aligned} \quad (12)$$

式中:  $\alpha \in (0.5, 1)$ ;  $S(T_n)$  是节点  $T_n$  的 treesize 属性;  $I(T_n)$  表示子树上无效节点数量。

$\alpha_{bal}$  和  $\alpha_{del}$  定义为:

$$\begin{aligned} \alpha_{bal}(T_n) &= \frac{\max\{S(T_n.leftson), S(T_n.rightson)\}}{S(T_n) - 1} \\ \alpha_{del}(T_n) &= \frac{I(T_n)}{S(T_n)} \end{aligned} \quad (13)$$

如果 IK-D Tree 不满足任一条件, 则需对 IK-D Tree 进行双线程重建。

(4) 在完成 IK-D Tree 构建后利用节点的“range”

信息加速临近点搜索。

### 2.2.2 ICP 算法精配准

查找到最近点后, 采用奇异值分解算法求取位姿增量, 获取其相对坐标转换关系: 旋转矩阵  $R_i$  与平移矩阵  $T_i$ 。通过迭代的方式求解使得公式中目标函数达到最小的刚体变换, 并将该变换作用于源点云, 迭代进行这一过程直到满足某一设定的收敛准则。

$$f(R_i, T_i) = \sum_{i=1}^N \|b_i - (R_i \cdot a_i + T_i)\|^2 \rightarrow \min \quad (14)$$

式中:  $a_i$  与  $b_i$  分别代表两组扫描数据中各点的三维坐标。当配准结果相对于真实值三轴位姿精度小于  $1^\circ$  时迭代终止。

通过以上配准过程获得最优变换矩阵, 完成六维位姿估计。位置估计可由平移向量直接获得, 姿态估计由旋转矩阵间接得到。设最优变换矩阵为:

$$M_2 = \begin{bmatrix} R_i & T_i \\ 0 & 1 \end{bmatrix} \quad (15)$$

则目标从物体坐标系到雷达坐标系的转换关系为:

$$\begin{bmatrix} X_L \\ Y_L \\ Z_L \\ 1 \end{bmatrix} = M_2 \cdot M_0^{-1} \cdot M_1 \cdot \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} \quad (16)$$

式中:  $M_2$  为 ICP 配准得到的位姿变换矩阵;  $M_0$  为雷达坐标系到彩色相机坐标系的外参矩阵;  $M_1$  为 PnP 求得的位姿变换矩阵。令:

$$M_4 = M_2 \cdot M_0^{-1} \cdot M_1 = \begin{bmatrix} R_4 & T_4 \\ 0 & 1 \end{bmatrix} \quad (17)$$

其中, 旋转矩阵为:  $R_4 = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$  平移向量

$T_4 = [t_1 \ t_2 \ t_3]^T$ 。传感器 6 自由度位姿估计为  $\Psi=(x,y,z,\alpha,\beta,\gamma)$ , 则位置与平移向量的关系为:

$$(x, y, z) = (t_1, t_2, t_3)$$

姿态与旋转矩阵的关系:

$$\begin{cases} \alpha = \arctan(r_{32}/r_{33}) \\ \beta = \arcsin(-r_{13}) \\ \gamma = \arctan(r_{12}/r_{11}) \end{cases} \quad (18)$$

## 3 实验验证

### 3.1 实验环境搭建

在仿真实验中, 所需仿真数据为仿真点云数据和仿真图像数据。实物模型如图 10(a) 所示, 尺寸约为

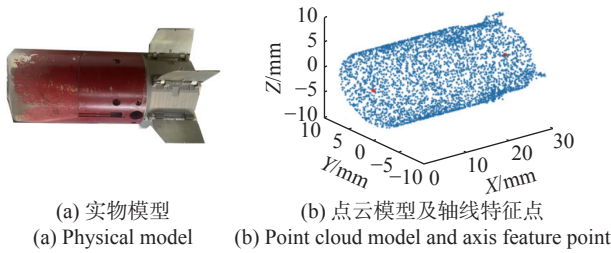


图 10 实物及点云模型

Fig.10 Physical object and point cloud model

500 mm×170 mm×170 mm。根据目标真实尺寸采用相关软件进行三维点云模型创建,并通过在点云模型中

增加轴线特征点用于后续运算。点云实物模型和点云模型如图 10(b) 所示,其中蓝色点为点云模型,红色点为轴线特征点。

仿真点云数据通过本实验室提出的三维点云模拟系统,模拟面阵雷达扫描模式采用双切片法<sup>[18]</sup>提取。假设模型初始位姿如图 11 所示,所采用的面阵激光雷达和彩色相机为实验室已有设备,参数信息如表 1 所示。利用 1.1 节中所提算法对雷达外参信息和彩色相机内参信息进行标定,仿真图像数据通过公式(4)将点云数据转换到图像像素坐标得到。

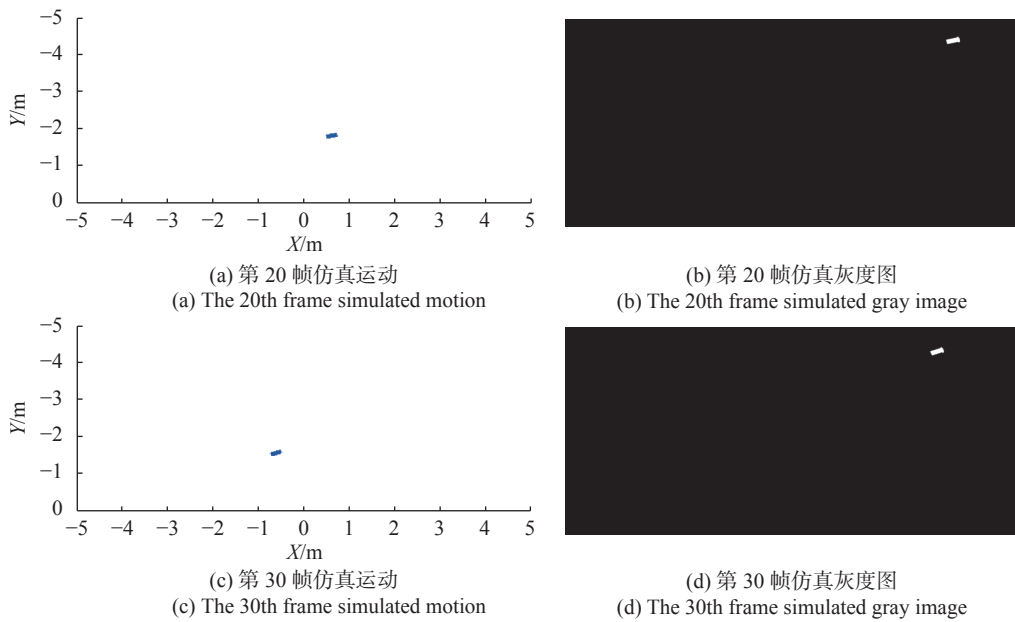


图 11 仿真数据

Fig.11 Simulation data

表 1 RS-LD1605 M 雷达参数范围

Tab.1 Parameter range of RS-LD1605 M radar

| Name                            | Value       |
|---------------------------------|-------------|
| Model                           | RS-LD1605 M |
| Horizontal field of view/(°)    | 56          |
| Vertical field of view/(°)      | 31          |
| Color image resolution          | 2448×1378   |
| Depth image resolution          | 640×360     |
| Horizontal angle resolution/(°) | 0.088       |
| Vertical angle resolution/(°)   | 0.088       |
| Detection distance/m            | 50          |
| Detection accuracy/cm           | 5           |
| Frame rate/Hz                   | 12          |

目标运动模型:飞行轨迹满足参考文献 [19]。以目标几何中心建立物体坐标系,绕 Y 轴旋转角度  $\beta$  为  $0.64(^{\circ})/\text{帧}$ ,绕 Z 轴旋转角度  $\gamma$  为  $0.51(^{\circ})/\text{帧}$ ,飞行目标做 X 轴初速度为 20 m/s, Y 轴方向自由落体的平抛运动。

位姿估计指标要求:位置误差  $\leq 0.01$  m; 旋转角度误差  $\leq 5^{\circ}$ ; 解算时间  $\leq 2$  s/幅。

### 3.2 扫描成像实验

采用实验室自研软件建立仿真环境,按照雷达成像机理模拟激光雷达扫描成像,目标与激光雷达初始物距为 30 m。根据可见点判别方法完成对飞行目标进行多视角可测部位点云数据获取。其中部分仿真结果(第 20 帧、第 30 帧)如图 11 所示。由面阵激光雷达成像原理<sup>[20]</sup>可知,在仿真实验的运动模型下飞行

目标在运动过程中可看作静止状态。实验结果表明扫描成像数据输出率可达 1 s/帧。(文中算法结果 CPU AMD Ryzen 5, 在内存为 16 GB 的 PC 机上运行所得, 操作系统为 Windows 10)。

### 3.3 仿真试验结果分析

为检测算法的稳定性, 进行 50 次独立测试对位姿估计结果进行评估。其中部分结果(第 20 帧、第 30 帧)如图 12 和图 13 所示。

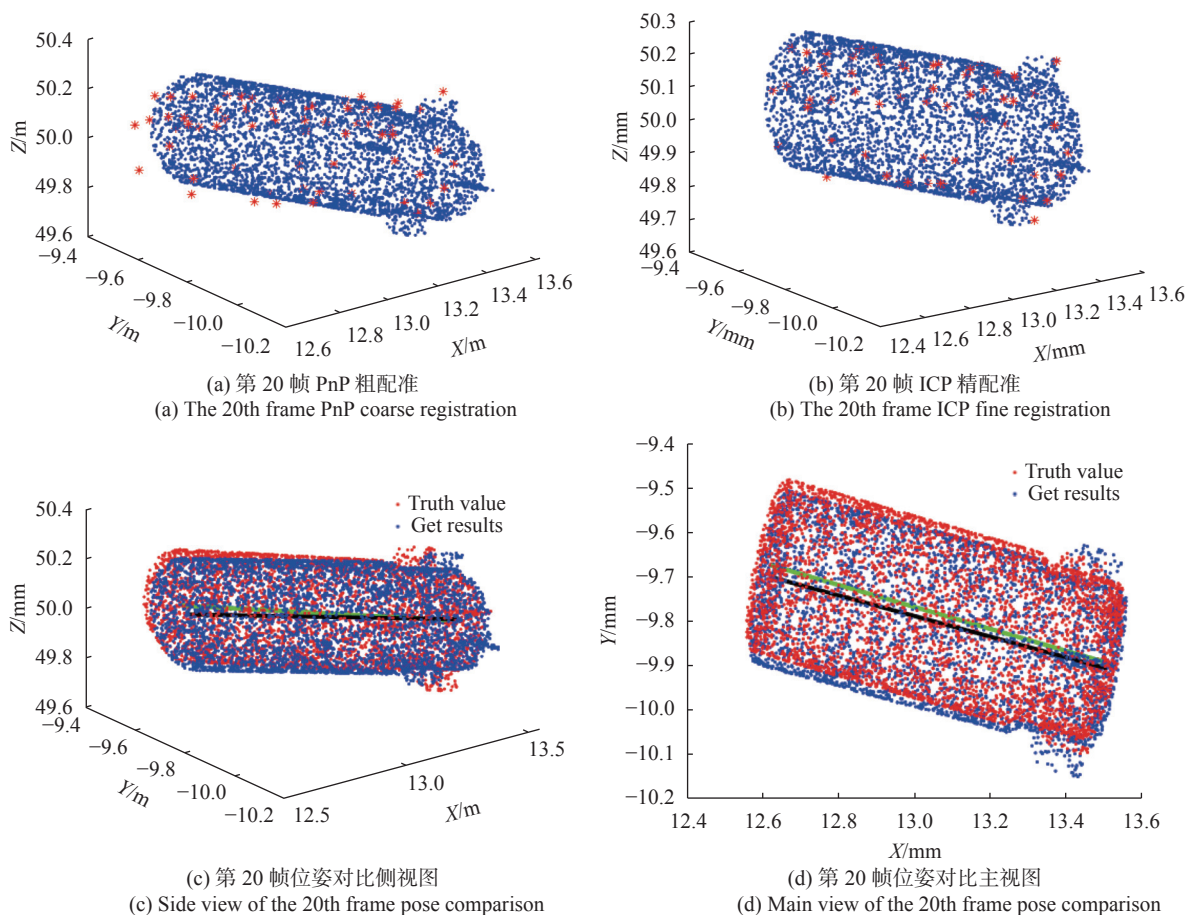
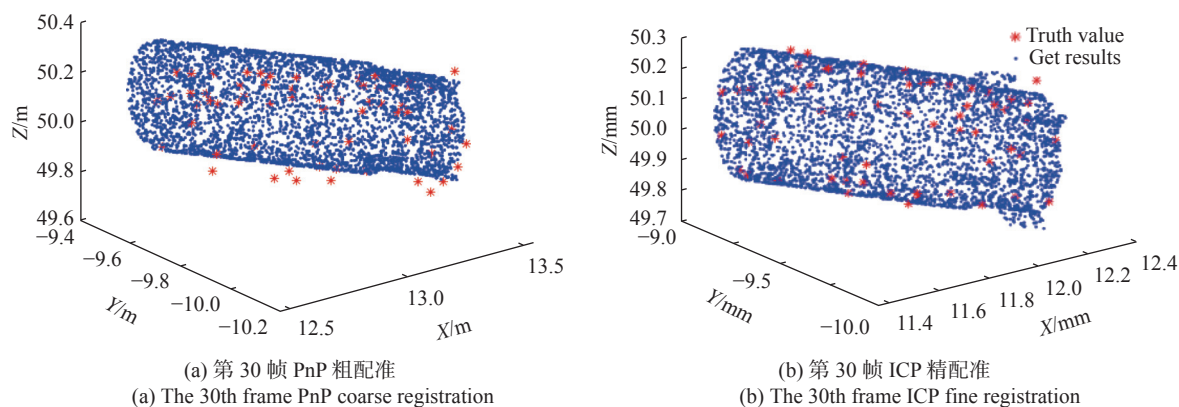


图 12 第 20 帧位姿获取过程展示

Fig.12 Display of pose calculation process in frame 20





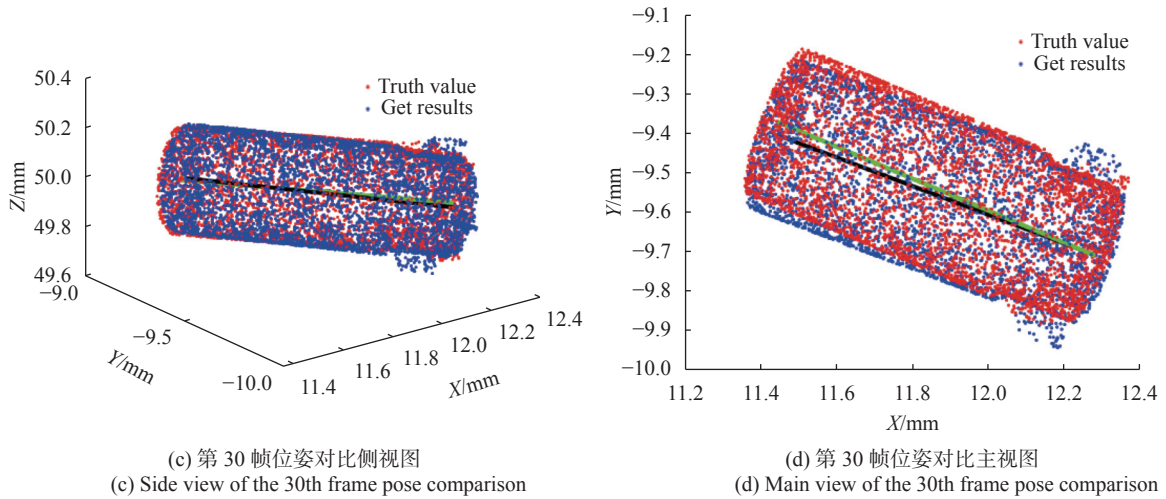


图 13 第 30 帧位姿求取过程展示

Fig.13 Display of pose calculation process in frame 30

解算得到的目标点云数据的轴线位置由轴线特征点可知,目标点云数据由 3.1 节中所提目标运动模型计算可得,然后再由轴线特征点可知目标的轴线位置。通过观察仿真目标点云数据与解算得到的目标点云数据的轴线与 Y、Z 轴的夹角来判断旋转角度误差。

目标位姿数据包括位置参数 ( $P_x, P_y, P_z$ ) 和姿态参数 ( $\alpha, \beta, \gamma$ ), 其中  $P_x, P_y, P_z$  为 X、Y、Z 轴的位置参数,

$\alpha$  为绕 X 轴旋转角度,  $\beta$  为绕 Y 轴旋转角度,  $\gamma$  为绕 Z 轴旋转角度。由于目标为类圆柱形, 因此  $\alpha$  不在此次实验结果的参考值。定义估计误差为 ( $\Delta P_x, \Delta P_y, \Delta P_z, \Delta \beta, \Delta \gamma$ )。对三轴平移距离误差及旋转角度误差、算法过程耗时进行了统计分析, 平均误差如图 14 所示。为证明文中算法的高精度性, 采用 ICP 算法对仿真数据进行直接配准, 误差结果如图 15 所示。与参考文献 [9-10] 所提算法进行对比, 比较效果见表 2。

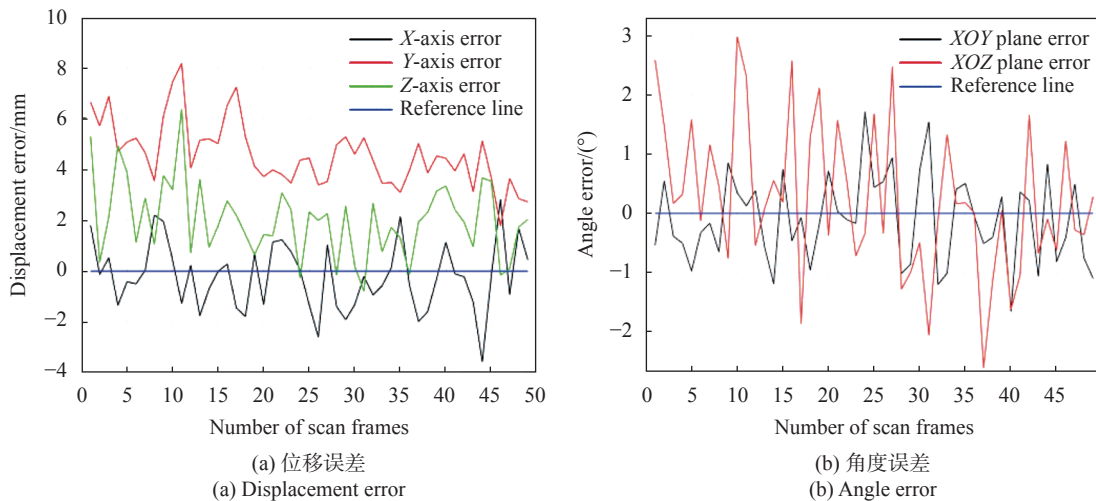


图 14 位姿求解误差

Fig.14 Pose solution error

由表 2 可知,三轴平移误差低于 ICP 算法和其他算法; 文中算法采用 PnP 算法完成点云粗配准, 获取点云间初始旋转平移矩阵; 利用 IK-D Tree 加速临近

点搜索, 采用 ICP 算法完成点云精配准。因此运算速度优于 ICP 算法劣于 PnP 算法。

实验结果表明, 在不同运动姿态下文中算法均可

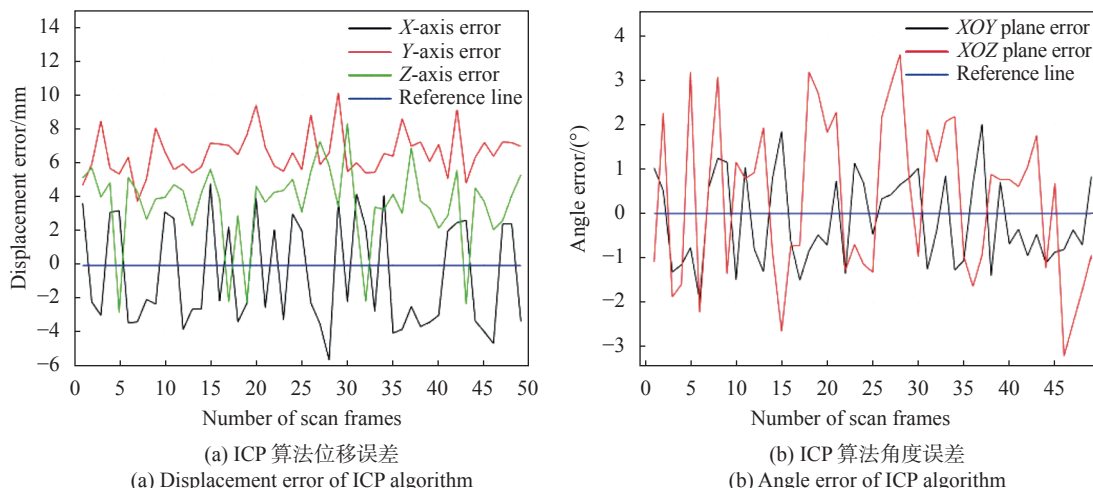


图 15 ICP 算法误差

Fig.15 ICP algorithm error

表 2 算法精度对比

Tab.2 Comparison of algorithm accuracy

| Group                 | $\Delta P_x/mm$ | $\Delta P_y/mm$ | $\Delta P_z/mm$ | $\Delta\beta/(^\circ)$ | $\Delta\gamma/(^\circ)$ | Time/ms |
|-----------------------|-----------------|-----------------|-----------------|------------------------|-------------------------|---------|
| Proposed algorithm    | 1.06            | 4.59            | 2.07            | 0.63                   | 1.01                    | 132     |
| ICP algorithm         | 2.93            | 8.37            | 4.21            | 0.94                   | 1.62                    | 261     |
| PnP algorithm         | 1.65            | 5.45            | 4.11            | 0.76                   | 1.33                    | 37      |
| Algorithm in Ref.[10] | 2.51            | 6.02            | 3.88            | 0.82                   | 1.41                    | 327     |

完成目标位姿估计。所提方法具有较高的估计精度，同时也证明了所提方法的有效性和准确性。

### 3.4 半物理仿真试验及结果分析

#### 3.4.1 建立仿真环境

建立半实物仿真试验环境对算法有效性进行验证。由于实验室不具备对炮弹实物进行抛洒的场地条件和抛洒装置条件。遂采用约为 3 : 1 缩放模型的铁质圆柱筒完成半物理实验。目标模型长宽高为 160 mm×50 mm×50 mm，激光雷达和彩色相机参数信

息与数字仿真试验一致。采用人工抛洒目标的方法模拟目标飞行过程进行了两组试验，目标距离雷达相对距离为 [10 m, 11 m]范围内变化，测试不同环境(A 组背景为静态建筑物和树木; B 组背景为花坛、草地和静态建筑物)、不同光照条件(A 组拍摄时间为上午 10 点, 天气状况: 晴; B 组拍摄时间为下午 4 点, 天气状况: 多云)下的文中算法的估计效果。每组试验条件下拍摄图像数量在 20 张以上, 部分拍摄图像如图 16 所示。



(a) 场景 A 彩色图  
(a) Scene A color picture



(b) 场景 B 彩色图  
(b) Scene B color picture

图 16 两场景彩色图片

Fig.16 Two scene color pictures

3.4.2 图像目标提取及性能评估

为检测图像目标检测算法的稳定性与有效性,对两组图片进行目标提取测试,并与背景差分法、传统 ViBe 算法以及参考文献 [5] 所提算法进行比对,对结

果进行统计与分析。目标提取结果如图 17 所示,为定量比较改进算法与原算法的性能,采用了正确率 (Precision, P)、召回率 (Recall, R)、错误分类比 (Percentage of Wrong Classifications, PWC) 作为评价指标,定义如下:

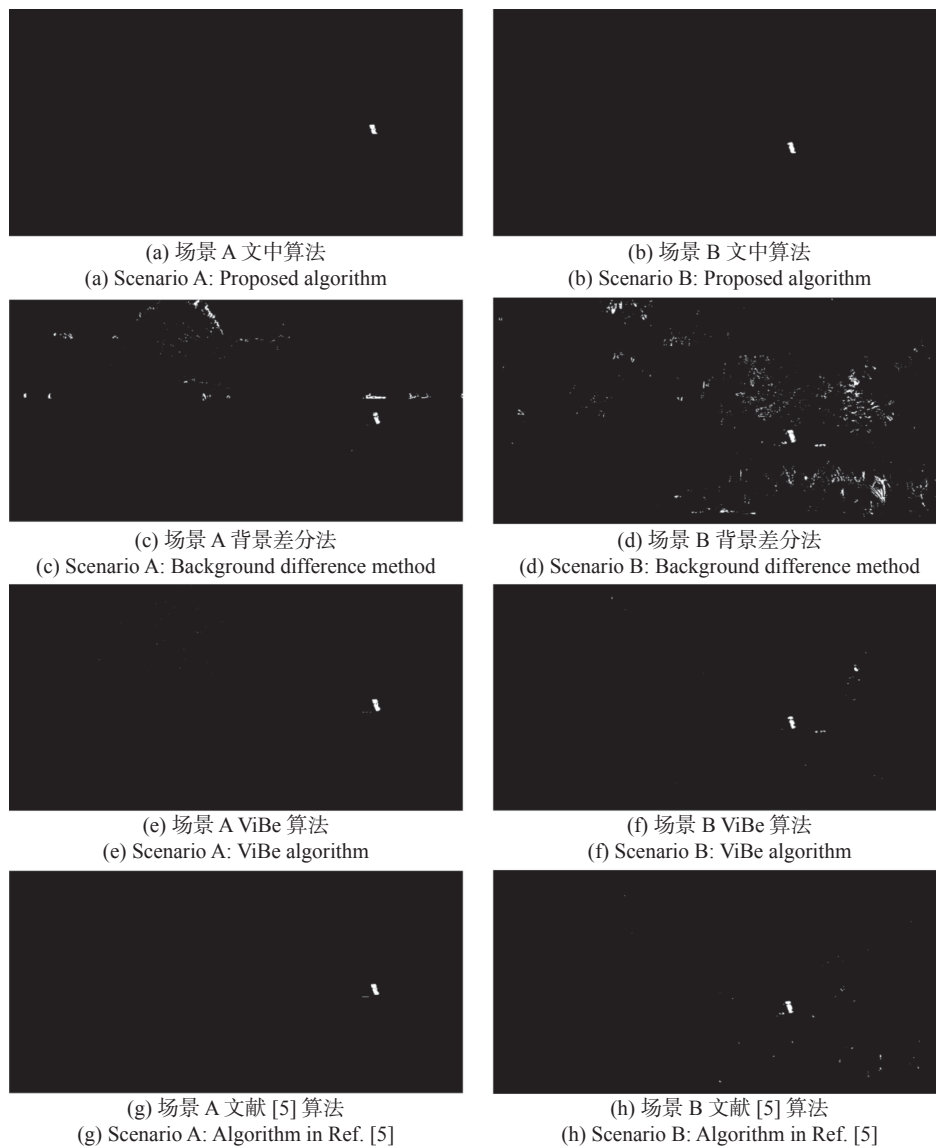


图 17 图像目标提取算法对比

Fig.17 Comparison of image target extraction algorithms

$$P = \frac{TP}{TP + FP} \quad (19)$$

$$R = \frac{TP}{TP + FN} \quad (20)$$

$$PWC = \frac{FP + FN}{FP + TN + FP + FN} \times 100\% \quad (21)$$

式中:  $TP$  为正确检测为前景像素的个数;  $FP$  为错误检测为前景像素的个数;  $TN$  为正确检测为背景像素

的个数;  $FN$  为错误检测为背景像素的个数;  $P$  为正确判定为前景像素的个数占判定为前景像素总数的百分比;  $R$  为正确判定为前景像素的个数占真正前景像素总数的百分比。提取点数以及性能评估见表 3~5。

由表 3~4 可见背景差分法在场景存在动态干扰物时提取效果不佳; 在背景为静态建筑物居多时, 参考文献 [5] 均有较好的前景目标提取效果, 但当背景

表 3 图像目标提取精度对比

Tab.3 Comparison of target extraction accuracy of image

| Scene   | Group                  | TP    | FP    | TN        | FN  | P     | R     | PWC      |
|---------|------------------------|-------|-------|-----------|-----|-------|-------|----------|
| Scene A | Our method             | 1 520 | 40    | 336 968 4 | 280 | 0.962 | 0.844 | 0.009 4% |
|         | Background subtraction | 1 627 | 6 015 | 336 372 9 | 173 | 0.213 | 0.904 | 0.183%   |
|         | Vibe algorithm         | 1 642 | 178   | 336 956 6 | 158 | 0.902 | 0.912 | 0.009 9% |
|         | Algorithm in Ref. [5]  | 1 732 | 117   | 336 962 7 | 68  | 0.936 | 0.962 | 0.005 5% |
| Scene B | Our method             | 1 607 | 31    | 336 931 3 | 393 | 0.978 | 0.803 | 0.013%   |
|         | Background subtraction | 1 627 | 7 302 | 336 204 2 | 373 | 0.182 | 0.813 | 0.228%   |
|         | Vibe algorithm         | 1 766 | 725   | 336 861 9 | 234 | 0.709 | 0.883 | 0.028%   |
|         | Algorithm in Ref. [5]  | 1 939 | 1 007 | 336 833 7 | 61  | 0.658 | 0.969 | 0.032%   |

表 4 图像目标提取耗时

Tab.4 Time consuming for image target extraction

| Group                  | A/ms | B/ms |
|------------------------|------|------|
| Proposed algorithm     | 301  | 289  |
| Background subtraction | 153  | 147  |
| Vibe algorithm         | 271  | 253  |
| Algorithm in Ref.[5]   | 206  | 198  |

为动态植物时,背景差分法和参考文献 [5] 提取效果不佳。ViBe 算法在静态背景和动态背景下均有较好的提取效果。但当光线不足且有变化时会产生部分噪点,文中算法召回率相比其他算法略有降低,但在正确率以及错误分类比上都有较好的提升,与传统 ViBe 算法相比正确率提高了 27%,错误分类比降低了 36%。背景差分法和参考文献 [5] 算法直接将运动图像和背景相比,运算量低鲁棒性差且两种方法均为手动设置阈值,因此与文中所提算法相比耗时短;文中算法相比与传统 ViBe 算法相比,将深度信息加入背景模型样本中作为运动目标检测的权值且采用自

表 5 目标位姿求取结果

Tab.5 Calculation results of target pose

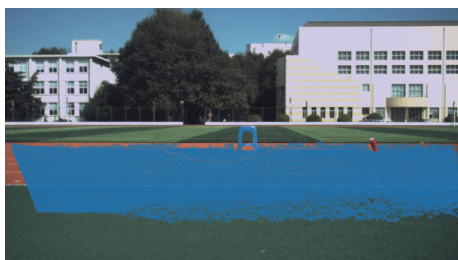
| Scene number | $\Delta P_x/mm$ | $\Delta P_y/mm$ | $\Delta P_z/mm$ | $\Delta\beta/(^\circ)$ | $\Delta\gamma/(^\circ)$ |
|--------------|-----------------|-----------------|-----------------|------------------------|-------------------------|
| Scene A      | 2.701 5         | -0.414 9        | 10.277 8        | 43.681 3               | 66.732 6                |
| Scene B      | 1.336 8         | 0.172 3         | 10.503 6        | 82.832 9               | 4.693 3                 |

适应阈值减少了人工干预,因此耗时有所提升。

3.4.3 位姿估计及性能评估

采用第 2 章中所提算法对目标点云进行提取和位姿估计,结果如图 18 和表 6 所示。

单次激光雷达扫描无法直接提取飞行目标的完整轮廓,从而无法掌握目标准确的位姿信息。因此通过彩色图像分别取目标的 4 个角点和目标每条边的中点作为特征点。利用 PnP 算法对飞行目标位姿进行两次估计,然后取两次估计结果的平均值与文中位姿求取结果进行对比,来分析实验的有效性。两次 PnP 算法的位姿求取结果如表 7 和表 8 所示。求取平均值后与表 6 所求数据进行对比,误差结果如表 9 所示。



(a) 场景 A 图像与点云融合  
(a) Scene A image and point cloud fusion



(b) 场景 B 图像与点云融合  
(b) Scene B image and point cloud fusion

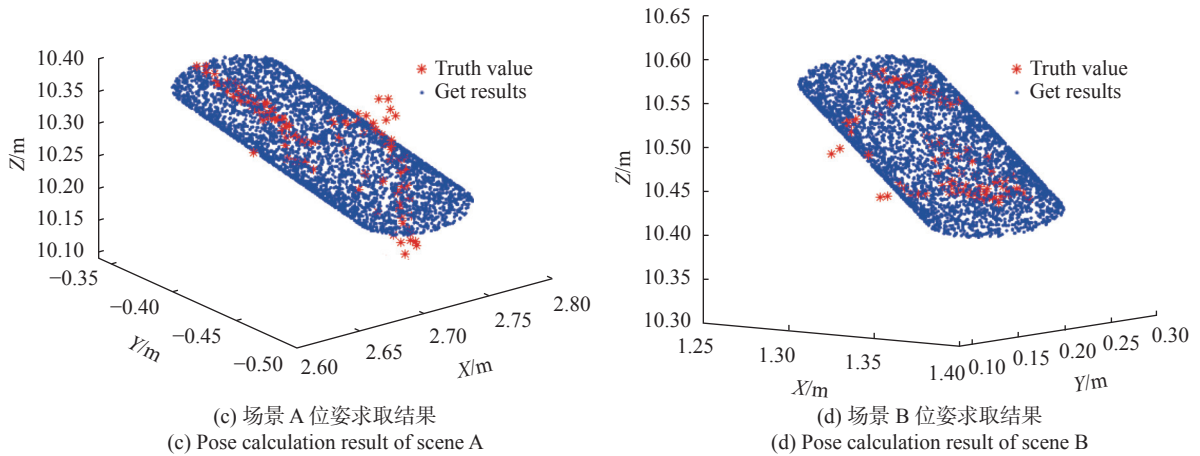


图 18 不同场景目标位姿求取

Fig.18 Calculation of target pose in different scenes

表 6 PnP 位姿求取

Tab.6 PnP pose calculation

| Scene number | $\Delta P_x/mm$ | $\Delta P_y/mm$ | $\Delta P_z/mm$ | $\Delta\beta/(^\circ)$ | $\Delta\gamma/(^\circ)$ |
|--------------|-----------------|-----------------|-----------------|------------------------|-------------------------|
| Scene A      | 2.7073          | -0.4213         | 10.2818         | 44.3518                | 65.3863                 |
| Scene B      | 1.3328          | 0.1685          | 10.5072         | 81.7567                | 5.3052                  |
| Scene A      | 2.7062          | -0.4137         | 10.2747         | 44.2721                | 65.8147                 |
| Scene B      | 1.3317          | 0.1705          | 10.5093         | 81.2602                | 5.3872                  |

表 7 位姿误差对比

Tab.7 Position and attitude error comparison

| Scene number | $\Delta P_x/mm$ | $\Delta P_y/mm$ | $\Delta P_z/mm$ | $\Delta\beta/(^\circ)$ | $\Delta\gamma/(^\circ)$ |
|--------------|-----------------|-----------------|-----------------|------------------------|-------------------------|
| Scene A      | 5.25            | 2.6             | 4.5             | 0.63                   | -1.13                   |
| Scene B      | -4.55           | -2.8            | 4.65            | -1.32                  | 0.65                    |

表 8 不同表面颜色下的点云采集率

Tab.8 Point cloud collection rate under different surface colors

| Colour           | White | Blue | Red  |
|------------------|-------|------|------|
| Acquisition rate | 1     | 0.97 | 0.94 |

表 9 不同物距和表面颜色下的点云数

Tab.9 Number of point clouds at different object distances and surface colors

| Group         | $\Delta P_x/mm$ | $\Delta P_y/mm$ | $\Delta P_z/mm$ | $\Delta\beta/(^\circ)$ | $\Delta\gamma/(^\circ)$ |
|---------------|-----------------|-----------------|-----------------|------------------------|-------------------------|
| Group A       | 0.97            | 4.26            | 1.84            | 0.58                   | 0.97                    |
| Group B       | 0.061           | 0.062           | 0.039           | 0.014                  | 0.082                   |
| Our algorithm | 1.06            | 4.59            | 2.07            | 0.63                   | 1.01                    |

半实物仿真试验结果显示,位姿误差小于指标要求。文中算法可以正确有效的对于目标点进行提取和位置求取,并且能够稳定适用于实际环境之中。

### 3.5 测量精度分析

#### 3.5.1 环境因素引起的测量误差

激光雷达对目标进行扫描时,目标物与激光雷达的距离和采集点云数呈负相关;目标表面颜色对脉冲信号的反射强度影响较大。点云反射强度从高到低依次为白色、绿色、蓝色、红色、黑色<sup>[21]</sup>。目标的表面粗糙度对点云数量和点云反射强度均无明显影响。采用参考文献 [21] 中提供的目标颜色对点云影响结果,假设白色时目标物的采集率为 1,则不同颜色条件下对采集点云数量的影响如表 10 所示。在文中预设仿真条件下,分别在 30 m、15 m 和 5 m 处完成飞行目标点云采集并根据提取率在目标点云中取随机点的方式实现目标不同颜色下的取点模拟,统计不同颜色下采集点的数量,结果如表 11 所示。

表 10 位置误差

Tab.10 Position error

| Group      | $\Delta P_x/mm$ | $\Delta P_y/mm$ | $\Delta P_z/mm$ |
|------------|-----------------|-----------------|-----------------|
| Group A    | 0.97            | 4.26            | 1.84            |
| Group B    | 0.061           | 0.062           | 0.039           |
| Our method | 1.06            | 4.59            | 2.07            |

由表 11 可知,影响点云采集数量的主要因素为激光雷达距离点云的距离。虽然目标表面颜色也对点云点数有所影响,但由于相差的点云数量过少,在

表 11 角度误差  
Tab.11 Angle error

| Group      | $\Delta\beta(^{\circ})$ | $\Delta\gamma(^{\circ})$ |
|------------|-------------------------|--------------------------|
| Group A    | 0.58                    | 0.97                     |
| Group B    | 0.014                   | 0.082                    |
| Our method | 0.63                    | 1.01                     |

ICP 配准时可以忽略不计。

### 3.5.2 算法引起的测量误差

在文中设计算法中,采用改良 ViBe 算法对目标进行提取时会丢失一部分像素点,经图像形态学处理后可得到完整目标,在恢复完整目标时采用膨胀和腐蚀操作,会使得到的目标与真实目标尺寸略有差别;采用 LSD 直线检测算法和霍夫变换对图像特征点进行提取时会出现提取误差;采用改进的 ICP 算法对目标点云进行位姿估计,目标点云的数量影响配准的精度。

### 3.5.3 对比验证

针对以上测量误差原因,采用控制变量法设置对照组与文中试验结果进行对比验证。分别设置两个对照组与表 2 中文中算法造成误差进行比较。A 组实验条件为:物距 30 m,手动提取图像特征点;B 组实验条件为:物距 5 m,手动提取图像特征点;

由文中算法与 A 组对比可知,采用手动提取图像特征点可以减小误差,但效果不佳;由文中算法与 B 组对比可知,当激光雷达与目标的距离缩短时,可以有效减少误差。

## 4 结 论

文中提出一种基于双模态信息融合飞行目标位姿估计算法,选择合适参数的基础上,可有效估计飞行目标的位姿信息。首先使用改进的 ViBe 算法对图像目标进行提取,较好的克服了传统 ViBe 算法光照变化不敏感问题,提高了搜索率降低了错误分类比;然后利用点云和图像融合完成对点云目标的快速提取;最后采用 PnP 算法获取点云间初始旋转平移矩阵,并采用基于 IK-D Tree 的 ICP 算法完成目标位姿估计,加快最近点搜索并提高了位姿估计精度。弥补了单源数据难以全面描述运动目标的缺陷,为飞行目标位姿估计提供了一个客观的解决方案。可应用于

飞行目标落点精准预测和制导控制,具有较高军事应用价值。

### 参考文献:

- [1] Pan Chang. An improved algorithm of missile falling point prediction [J]. *Journal of Ordnance Equipment Engineering*, 2018, 39(9): 30-32+46. (in Chinese)
- [2] Liu Junliang, Chen Shangfeng, Lu Huanzhang, et al. Research on infrared irradiance data stimulation for ballistic targets recognition [J]. *Infrared and Laser Engineering*, 2016, 45(10): 39-45. (in Chinese)
- [3] Li Jiawei, Jiang Jing, Liu Chongyang, et al. Modeling and simulation for target complex trajectory of ballistic missile [J]. *Journal of System Simulation*, 2020, 32(8): 1515-1523. (in Chinese)
- [4] Shao Huibing, Cui Naigang, Wei Changzhu. Multi-constrained intelligent trajectory planning for gliding missiles [J]. *Guangxue Jingmi Gongcheng/Optics and Precision Engineering*, 2019, 27(2): 410-420. (in Chinese)
- [5] Le Ying, Zhao Zhicheng. Multi-moving object detection and segmentation based on background difference method [J]. *Chinese Journal of Construction Machinery*, 2020, 18(4): 305-309. (in Chinese)
- [6] Barnich O, van Droogenbroeck M. ViBe: a universal background subtraction algorithm for video sequences [J]. *IEEE Transactions on Image Processing*, 2011, 20(6): 1709-1724.
- [7] Liu Jiajun, Lin Haokun. Adaptive threshold moving target detection algorithm based on ViBe method [J]. *Laser & Optoelectronics Progress*, 2022, 59(2): 144-153. (in Chinese)
- [8] Hu Tao, Zhu Xinyan, Guo Wei, et al. A moving object detection method combining color and depth data [J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(2): 276-282. (in Chinese)
- [9] Yin Fang, Wu Yun. Model reconstruction and pose optimization of non-cooperative rotating space target [J]. *Guangxue Jingmi Gongcheng/Optics and Precision Engineering*, 2019, 27(8): 1854-1862. (in Chinese)
- [10] Tian Guangbao, Wang Jian, Wang Bowen. Monocular camera non-cooperative target extraction and pose detection [J]. *Infrared and Laser Engineering*, 2021, 50(12): 536-544. (in Chinese)
- [11] Liu yaohua, Ma yue, Xu min. Point cloud instance segmentation based robust multi-target [J]. *Laser & Optoelectronics Progress*, 2022, 20(8): 1-18. (in Chinese)
- [12] Kechagias Stamatis O, Aouf N, Dubanchet V, Richardson M. A.

- DeepLO: Multi-projection deep LIDAR odometry for space orbital robotics rendezvous relative navigation [J]. *Acta Astronautica*, 2020, 177: 270-285.
- [13] Tan Changsheng, Huang Genghua, Wang Fengxiang et al. Optimization and Validation of Coherent Point Drift for Planar-array- Based Point Cloud in Space Pose Measuremn[J]. *Infrared and Laser Engineering*, 1-12[2022-09-19]. (in Chinese)
- [14] Bi Yong, Pan Mingqi, Zhang Shuo, et al. Overview of 3 D point cloud super-resolution technology [J]. *Chinese Optics*, 2022, 15(2): 210-223. (in Chinese)
- [15] Li Ronghua, Wen Shuaifang, Xiao Yuzhi, et al. Indexes optimization macro model of on-orbit imaging lidar in space [J]. *Infrared and Laser Engineering*, 2018, 280(2): 164-171. (in Chinese)
- [16] Lv Yuanzhi, Sun Qiang, Bi Guoling. Recognition and repairing of surface hole in three dimensional laser scanning system [J]. *Chinese Optics*, 2016, 9(1): 114-121. (in Chinese)
- [17] Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing [J]. arXiv preprint arXiv: 2107.13586, 2021.
- [18] Li Ronghua, Xue Haopeng, Yang Youqing, et al. Linear array lidar imaging mechanism and 3 D reconstruction for space instable targets [J]. *Journal of Astronautics*, 2020, 41(8): 1049-1057. (in Chinese)
- [19] Xu Fei, Xu Shimeng, Du Jianhua, et al. Research on trajectory correction of equipment with autonomous precise fixed-point drop [J]. *Journal of Ordnance Equipment Engineering*, 2017, 38(6): 165-168. (in Chinese)
- [20] Fang Yi, Zhang Xiuda, Hu Jian, et al. A method applying gray image superposition to improve ranging accuracy in planar array laser radar [J]. *Chinese Journal of Lasers*, 2013, 40(1): 0114002-230. (in Chinese)
- [21] Cai Yue, Xu Wenbing, Liang Dan, et al. Impacts of different factors on accuracy of point cloud obtained from terrestrial three-dimensional laser scanning [J]. *Laser & Optoelectronics Progress*, 2017, 54(9): 092801-373. (in Chinese)

## Pose estimation of flying target based on bi-modal information fusion

Li Ronghua<sup>1</sup>, Wang Meng<sup>1</sup>, Zhou Wei<sup>2</sup>, Fu Jiaru<sup>1</sup>

(1. Institute of Mechanical Engineering, Dalian Jiaotong University, Dalian 116028, China;

2. No.91550 Unit of the PLA, Dalian 116023, China)

### Abstract:

**Objective** Flying target pose estimation is the key technology to realize trajectory prediction and missile guidance control. Real-time calculation of missile posture is conducive to judge whether the missile hits the target, timely detect missile failure, and carry out early destruction. The development of information and intelligent technology has led to the improvement of data acquisition accuracy of color camera, laser radar and other sensors, which has formed a technical system of data acquisition by sensors and target position and attitude estimation by relevant algorithms. Most of the existing target pose estimation methods can effectively detect and estimate the target pose. However, there are some problems in the precision prediction and guidance control of missile landing point, such as the inability to quickly and accurately extract and estimate the position and attitude of the flying target in the complex background. Therefore, on the premise of ensuring being real-time, a method for estimating the pose of flying targets based on area array lidar and bi-modal information fusion is proposed.

**Methods** Firstly, the coordinate transformation model between the camera and the lidar is established to realize the pixel level matching of the two sensors, fusing the image and point cloud at the same time (Fig.2); Secondly, the ViBe (Visual Background Extractor) algorithm and depth information fusion algorithm are used to extract the moving target in the image, and select the corresponding point cloud according to the image moving target position box (Fig.5); Finally, the PnP (Perspective-n-Point) algorithm is used for rough registration of feature points (Fig.8), to obtain the initial rotation and translation matrix between point clouds. And using I-Kd Tree (Incremental K-dimensional Tree) to accelerate the search of adjacent points, the ICP (Iterative Close Point)

algorithm is used for accurate registration, to improve the registration speed.

**Results and Discussions** Simulation test and hardware-object simulation test are used to verify the accuracy and stability of the method. The results show that the accuracy of the two-dimensional image object detection algorithm is 97% (Tab.3), and the error classification ratio is 0.0112% (Tab.3). Compared with the traditional ICP algorithm, the accuracy of the pose estimation algorithm is improved by 53.2% (Tab.2), the single time consumption is reduced to 132 ms from 261 ms (Tab.2). Compared with other algorithms, the pose estimation algorithm also has certain advantages.

**Conclusions** An algorithm for estimating the pose of flying targets based on bi-modal information fusion is proposed, which can effectively estimate the pose of flying targets on the basis of selecting appropriate parameters. The accuracy of the algorithm is verified by simulation tests. 50 frames of data are simulated and the average error is calculated under the initial condition that the initial object distance between the target and the lidar is 30 m. The simulation results show that the  $X$ -axis error is 1.06 mm, the  $Y$ -axis error is 4.59 mm, the  $Z$ -axis error is 2.07 mm, the  $Y$ -axis rotation angle error is  $0.63^\circ$ , the  $Z$ -axis rotation angle error is  $1.01^\circ$ , and the solution time is 132 ms. The accuracy of the algorithm is verified by semi-physical ground experiments, and the precision ( $P$ ) in the image target extraction test is 0.97, the recall ( $R$ ) is 0.844, and the percentage of wrong classification (PWC) is 0.011 2%. The statistical average error in the pose estimation test is respectively 4.9 mm for  $X$ -axis error, 2.7 mm for  $Y$ -axis error, 4.62 mm for  $Z$ -axis error,  $0.97^\circ$  for  $Y$ -axis rotation angle error and  $0.89^\circ$  for  $Z$ -axis rotation angle error. The defect that the single source data of the proposed method is difficult to describe the moving target comprehensively can be remedied, and an objective solution is provided for the position and attitude estimation of the flying target. The proposed method is applied to the accurate prediction and guidance control of the landing point of flying targets, and has high military application value.

**Key words:** bi-modal information fusion; pose estimation; flying target; target extraction

**Funding projects:** Young Scientists Fund of the National Natural Science Foundation of China (41901415); Science and Technology Foundation of State Key Laboratory Grant Number (2022-JCJQ-L8-015-0201); The Key Project of Scientific Research Project of Education Department of Liaoning Province (LJKZ0475); Dalian High-Level Talent Innovation Support Program Grant Number (2022RJ03)