

文章编号 1004-924X(2015)06-1722-10

基于多直线对应和加权最小二乘的位姿估计

张跃强^{1,2*}, 苏 昂^{1,2}, 刘海波^{1,2}, 尚 洋^{1,2}, 于起峰^{1,2}

(国防科技大学 航天科学与工程学院, 图像测量与视觉导航湖南省重点实验室, 湖南 长沙 410073)

摘要: 为了求解复杂环境下目标的位姿, 提出了基于多直线对应的加权最小二乘位姿估计算法。首先对模型直线进行等间隔采样, 并沿采样点投影法线方向搜索图像点对应; 然后利用图像点对应局部和全局特性对样本点进行加权; 最后通过优化法向距离实现目标位姿的优化求解。为了解决模型-图像对应错误引起的优化失败问题, 算法在模型-图像点匹配阶段为每个采样点保留多个图像点对应, 通过随机 Hough 变换(RHT)算法将图像点对应约束在直线上, 并为每条模型直线保留多图像直线对应。在对样本点进行加权时, 综合考虑了样本点自身的属性和样本点同周围点的关系, 有效提高了算法对纹理, 背景, 噪声等的鲁棒性。实验结果表明: 提出的方法能够实现复杂环境下目标位姿的优化求解, 其在 x 方向、 y 方向和 z 方向的角度估计误差分别优于 0.4 、 0.3 和 0.1° ; 在垂直光轴方向和沿光轴方向的相对位置误差则分别优于 0.03% 和 0.1% 。相比单假设方法, 提出的方法能够更有效地克服复杂背景干扰, 实现特殊视图目标位姿的稳定估计。

关键词: 机器视觉; 位姿估计; 三维跟踪; 多假设; 迭代加权最小二乘

中图分类号: TP391.4, TP242.6 **文献标识码:** A **doi:** 10.3788/OPE.20152306.1722

Pose estimation based on multiple line hypothesis and iteratively reweighted least squares

ZHANG Yue-qiang^{1,2*}, SU Ang^{1,2}, LIU Hai-bo^{1,2}, SHANG Yang^{1,2}, YU Qi-feng^{1,2}

(Hunan Provincial Key Laboratory of Image Measurement and Vision Navigation,
College of Aerospace Science and Engineering,
National University of Defense Technology, Changsha 410073, China)

* Corresponding author, E-mail: zyoungnuds@yahoo.com

Abstract: To estimate the pose of known rigid objects efficiently in a complex environment, a rigid object pose estimation method was proposed by combining multiple line hypothesis and iteratively reweighted least squares. The 1D search was utilized to obtain the corresponding image point along a normal direction for each model sample point by an equal interval sampling. Then, the weight of a sample point was calculated according to the local and global appearances of the corresponding image point for each visible model sample point. The optimized pose parameters were obtained by minimizing the errors between the sample points and their corresponding image points. To avoid the failure of the pose optimization caused by the mismatches of the model and image lines, multiple low level hypotheses were retained for each model sample point in the registration process and they were

收稿日期: 2015-01-19; 修订日期: 2015-03-06.

基金项目: 国家 973 重点基础研究发展计划资助项目 (No. 2013CB733100); 国家自然科学基金资助项目 (No. 11332012)

classified into multiple lines for each potential edge by the Random Hough Transform(RHT). Due to the use of the property of the sample point as well as the relation to the neighbor points, the robustness to disordered background and noise was enhanced in the weighting process. Experiments show the proposed method effectively estimates the poses of freely moving objects in an unconstrained environment. The precisions of the poses on x , y and z axes are better than 0.4° , 0.3° and 0.1° respectively; and those of relative positions perpendicular to the optical axis and along the optical axis are better than 0.03% and 0.1% respectively. Comparisons with the single hypothesis based method demonstrate that the proposed method overcomes the influence induced by the complex background and optimizes the pose parameters in special views.

Key words: machine vision; pose estimation; 3D tracking; multiple hypothesis; Iteratively Reweighted Least Square(IRLS)

1 引言

位姿估计的目的是恢复出相机相对于场景的位置和旋转,或者目标相对于相机的相对运动。它广泛应用于增强现实、视觉伺服、机器人操作以及空间交会对接等领域。本文的研究主要面向航天器交会对接,其目标位姿对应于目标体坐标系相对于相机坐标系(同追踪器固联)的刚体变换参数。

传统的目标位姿估计方法一般需要在目标表面布置合作标志(标志点,标志线,平面编码标志等),这些标志在目标体坐标系的坐标精确已知。由于合作标志的成像模式较为显著,利用简单的图像处理手段便能实现它们的高精度提取和快速跟踪,且易于建立二维图像特征同三维合作标志的对应关系,然后通过 PnP(n 点透视)算法^[1-2]或者 PnL(n 直线透视)算法^[3-4]实现目标位姿的线性解算。在良好的实验条件下,该类方法拥有较高的鲁棒性和精度,且实时性较好。然而在实际情况下,由于合作标志容易被遮挡而不可见,或者因为成像条件的变化出现定位偏差,因此极易导致目标位姿解算失败或者精度降低。

对于大部分人造目标,尽管无法提供合作标志信息,但是目标自身的模型信息已知,充分利用目标模型的先验信息也可以实现目标位姿的解算。常用的模型信息有特征点,纹理以及直线等特征。本文研究的对象为自由运动的人造目标,该类目标纹理较为简单或者重复分布,利用点特征或者纹理特征很难实现目标位姿的跟踪。尽管人造目标缺少纹理,但其具有大量能够很好反应

目标几何信息的直线,相对于点特征,它们的提取更可靠,属性更为丰富。因此,本文利用目标的直线模型来实现目标位姿的估计。基于直线模型的目标位姿估计方法可以分为两类:需要显式提取直线的方法和无需显示提取直线的方法。

需要显式提取直线的方法^[5-6]首先在图像上提取直线,再将其与模型直线进行匹配,最后利用对应关系进行位姿估计和跟踪。该类方法的优点为可以快速计算模型直线同图像直线的距离(角度,位置等),实现快速 3D 跟踪;建立模型直线同图像直线的对应关系,利用基于直线的位姿解算方法可以实现目标位姿的优化求解。另外通过对同名直线的跟踪,可以利用多视图实现对目标关键结构模型的重建,当目标模型未知时,便可以将重建后的模型作为目标模型进行 3D 跟踪。该方法受制于直线提取结果不稳定以及复杂度较大等缺点。另外由于该类方法只适用于多面体目标,故在一定程度上限制了应用范围。

为了有效避免直线提取带来的测量不确定性,出现了很多无需显式提取直线的方法。该类方法多沿模型采样点法线方向寻找亮度或者梯度极大值点,并通过优化法向距离实现位姿的估计。该类方法无需在图像中提取直线,因而适应性较好,比较典型的为 RAPID 方法^[7],本文方法属于此类方法。为了提高跟踪稳定性,在基本的跟踪框架下,出现了三种方法:一是将直线特征同其他特征融合,如结合直线与关键点特征的方法^[8-9],结合直线和纹理特征的方法^[10-11],结合直线和颜色信息的方法^[12-13]。二是使用鲁棒估计方法^[14-15](如 M-估计、RANSAC)有效消除误匹配的影响,这种方法在图像测量阶段保留多个假设,

在模型采样点同图像点配准阶段通过鲁棒估计方法剔除野值。由于该算法只根据对应图像点计算一个姿态,因此错误的直线匹配仍然可能带来跟踪误差或者导致跟踪失败。三是使用非线性滤波法,如基于粒子滤波的方法^[16-17]。这种方法使用一定数目的加权粒子来逼近系统的后验概率密度,其中每个粒子都表示目标的一个可能位姿,由于该方法为目标保留了多个可能的位姿假设,因此鲁棒性更强,其缺点是算法复杂度较高。

为了提高杂乱背景干扰或者特殊视图条件下目标位姿优化的鲁棒性,本文提出了一种结合多假设直线对应和加权最小二乘的目标位姿优化算法。首先在模型图像配准阶段,为每个模型采样点保留了多个对应图像点,然后将图像对应点聚类为多条直线,并为每条模型直线保留多条图像直线对应;接着综合利用模型采样点自身的属性(方向距离,梯度模值)和样本点同周围点的关系(到对应拟合直线的距离以及所在直线的权重等)对模型采样点进行加权;最后利用加权最小二乘算法通过优化法向距离实现目标位姿的解算。由于在解算阶段为每条模型直线保留了多假设对应,因此本文方法有效克服了杂乱背景的影响,提高了特殊视图中优化的稳定性。

2 最小二乘算法基本原理

如图 1 所示,当前目标位姿条件下,模型直线的投影直线使用实线表示,图像直线使用虚线表示。

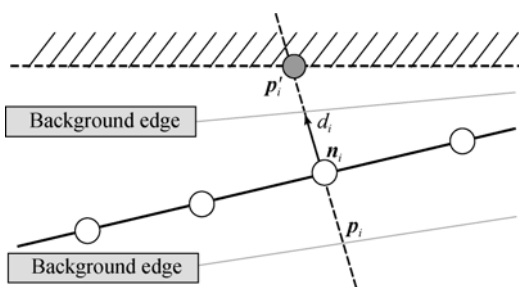


图 1 从模型直线到图像直线 1D 搜索配准示意图

Fig. 1 Searching schematic of 1D matching from model line to image line

对于模型直线上的每一个采样点 p_i ,沿其在直线的法线方向进行搜索,可以得到其对应的

图像边缘点 p'_i ,该边缘点到模型直线采样点的距离 d_i 称为法向距离:

$$d_i = \mathbf{n}_i^T (\mathbf{p}'_i - \mathbf{p}_i), \quad (1)$$

其中 \mathbf{n}_i 为模型直线采样点在当前位姿条件下的单位法向量。

假设相机成像模型为针孔模型,则目标体坐标系的点 \mathbf{P}^M 与其在图像坐标系的投影点 \mathbf{p} 关系如下:

$$\mathbf{p} = \Gamma(\mathbf{P}^M, \mathbf{E}) = \left[f_x \frac{(\mathbf{E}\mathbf{P}^M)_x}{(\mathbf{E}\mathbf{P}^M)_z} \quad f_y \frac{(\mathbf{E}\mathbf{P}^M)_y}{(\mathbf{E}\mathbf{P}^M)_z} \right]^T, \quad (2)$$

其中: $\mathbf{P}^C = \mathbf{E}\mathbf{P}^M$ 是目标体坐标系点 \mathbf{P}^M 在相机坐标系下的坐标, f_x, f_y 为相机 x, y 方向的等效焦距, \mathbf{E} 为目标系到相机系的刚体变换矩阵(对应目标的位姿):

$$\mathbf{E} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix}, \quad (3)$$

其中: $\mathbf{R} = \mathbf{R}_z(A_z) \mathbf{R}_y(A_y) \mathbf{R}_x(A_x)$ 为目标系到相机系的旋转矩阵, A_x, A_y, A_z 分别为目标系绕相机系 x 轴, y 轴, z 轴旋转的欧拉角(沿坐标轴正方向顺时针旋转为正,本文旋转顺序为 $x-y-z$); $\mathbf{T} = [T_x \quad T_y \quad T_z]^T$ 为目标系到相机系的平移向量; $(A_x, A_y, A_z, T_x, T_y, T_z)$ 为本文待求目标的姿态参数。

假设模型直线可见的采样点数目为 N ,利用目标当前位姿将模型直线采样点投影到图像上,并沿其法线方向 1D 搜索其对应的图像点,则法向距离最小化问题可以描述为:

$$\hat{\boldsymbol{\mu}} = \operatorname{argmin}_{\boldsymbol{\mu}} \sum_{i=1}^N (d_i)^2 =$$

$$\operatorname{argmin}_{\boldsymbol{\mu}} \sum_{i=1}^N (\mathbf{n}_i^T (\mathbf{p}'_i - \Gamma(\mathbf{P}_i^M, \mathbf{E})))^2. \quad (4)$$

给定目标帧间的刚体运动为 \mathbf{M} ,则 t 时刻的位姿 \mathbf{E}_t 和 $t+1$ 时刻的位姿 \mathbf{E}_{t+1} 之间的关系为:

$$\mathbf{E}_{t+1} = \mathbf{E}_t \mathbf{M}. \quad (5)$$

目标帧间刚体运动 \mathbf{M} 使用李群^[18]方法可以表示为:

$$\mathbf{M} = \exp\left(\sum_{j=1}^6 \mu_j \mathbf{G}_j\right), \quad (6)$$

其中 $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6)$ 为刚体运动参数,对应于帧间速度和角速度,基矩阵的形式如下:

$$\mathbf{G}_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \mathbf{G}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{G}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \mathbf{G}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{G}_5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \mathbf{G}_6 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

令误差 d_i 对位姿参数的雅可比矩阵为 \mathbf{J} , 则 $\mathbf{J} \in \mathbf{R}^{N \times 6}$, 其中 N 为可见模型采样点的个数, 则 \mathbf{J} 的任一元素 J_{ij} 形式如下:

$$J_{ij} = \frac{\partial d_i}{\partial \mu_j} = \mathbf{n}_i^T \frac{\partial}{\partial \mu_j} (\Gamma(\mathbf{P}_i^M, \mathbf{E}_t \mathbf{M})) = \mathbf{n}_i^T \left\{ \frac{\partial \mathbf{p}_i}{\partial \mathbf{p}_i} \right\}_{2 \times 2} \left\{ \frac{\partial \tilde{\mathbf{p}}_i}{\partial \mathbf{P}_i^C} \right\}_{2 \times 4} \left\{ \frac{\partial \mathbf{P}_i^C}{\partial \mu_j} \right\}_{4 \times 1} = \mathbf{n}_i^T \mathbf{J}_K \mathbf{J}_M \tilde{\mathbf{J}}_{ij}, \quad (7)$$

其中:

$$\mathbf{p}_i = [f_x \tilde{u}_i \quad f_y \tilde{v}_i]^T,$$

$$\tilde{\mathbf{p}}_i = \begin{bmatrix} \tilde{u}_i \\ \tilde{v}_i \end{bmatrix} = \begin{bmatrix} x_i^C & y_i^C \\ z_i^C & z_i^C \end{bmatrix}^{-T},$$

$$\mathbf{P}_i^C = [x_i^C \quad y_i^C \quad z_i^C \quad 1]^T,$$

$$\mathbf{J}_K = \begin{bmatrix} f_x & 0 \\ 0 & f_y \end{bmatrix},$$

$$\mathbf{J}_M = \begin{bmatrix} \frac{1}{z_i^C} & 0 & \frac{-x_i^C}{(z_i^C)^2} & 0 \\ 0 & \frac{1}{z_i^C} & \frac{-y_i^C}{(z_i^C)^2} & 0 \end{bmatrix},$$

$$\tilde{\mathbf{J}}_{ij} = \mathbf{E}_i \mathbf{G}_j \mathbf{P}_i^M.$$

通过对透视投影模型进行线性化处理, 可以将式(4)描述的优化问题转化为最小二乘问题进行求解:

$$\mathbf{J} \boldsymbol{\mu} = \mathbf{d}, \quad (8)$$

其中 $\mathbf{d} = (d_1 \quad d_2 \quad d_3 \quad \dots \quad d_i \quad \dots \quad d_N)^T$ 为 3D 模型直线可见采样点投影到对应图像点的距离向量。

式(8)的最小二乘解为:

$$\hat{\boldsymbol{\mu}} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{d}. \quad (9)$$

求解得到相邻帧的运动后, 根据式(5)和式(6)可以得到目标 $t+1$ 时刻的位姿:

$$\mathbf{E}_{t+1} \approx \mathbf{E}_t \left(\mathbf{I} + \sum_{j=1}^6 \boldsymbol{\mu}_j \mathbf{G}_j \right). \quad (10)$$

得到目标 $t+1$ 时刻的位姿后, 根据式(3)可以得到目标系到相机系的旋转矩阵和平移向量,

并分解旋转矩阵得到欧拉角 A_x, A_y, A_z (详见参考文献[19])。至此给定目标位姿初值(上一时刻的位姿)通过最小化法向距离, 可以求解目标在当前时刻的位姿参数 $(A_x, A_y, A_z, T_x, T_y, T_z)$ 。

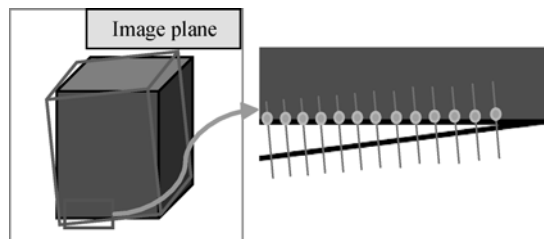
3 算法描述

当目标受杂乱背景或者噪声干扰时, 将出现大量的虚假直线, 需要利用加权最小二乘优化方法优化目标位姿。本文的加权分为全局和局部两部分。全局权值考虑了直线约束, 除了拟合误差还有模型直线同拟合直线的距离和夹角。局部权值用于衡量采样点对应的好坏的方法主要有法向距离、梯度模值以及梯度方向到图像边界的距离。

3.1 多假设直线对应

对于只包含目标、不含任何噪声的图像, 法向距离的测量误差可认为近似服从高斯分布, 此时最小二乘算法的精度很高。然而在实际图像中, 受噪声、阴影、复杂纹理或者杂乱背景的影响, 特别是当目标直线与背景干扰直线非常接近时, 法向搜索常常会得到多个候选边缘点和纹理边缘点。如果匹配点对应中存在野值点, 即使其数目很少, 利用最小二乘估计出的位姿结果也可能远远偏离真值, 从而导致求解失败。

另外在某些姿态下, 投影直线可能跟多目标强直线位置接近(此种条件下的目标视图称为特殊视图), 如图 2(b)所示, 此时沿法线方向也可能搜索到多个候选匹配点, 且这些候选匹配点都对应目标边缘点, 无法确定采样点与哪个边缘点对应, 选择任何一个对应都可能导致误匹配, 此种情况下最好能够保留 3D 目标模型直线的多条候选匹配直线。为此本文采用随机霍夫变换(RHT)自动剔除候选匹配点中的野值点, 并拟合多组模型参数, 为每条投影直线保留多条候选直线。



(a)一般视图

(a)General view

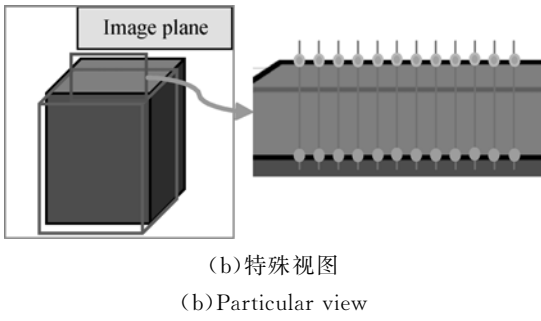


图 2 3D 模型投影直线对应图像直线

Fig. 2 Matching between 3D model projection lines and corresponding image lines

对于每条 3D 模型直线 $L_l (l \in [1, N_L], N_L$ 为可见的模型直线数目), 其采样点投影集合为 $\{\mathbf{p}_{l,i}\} (i \in [1, N_{l,i}], N_{l,i}$ 为模型直线 L_l 上的采样点个数), 对于模型直线上的采样点投影 $\mathbf{p}_{l,i}$, 沿其法向搜索梯度局部极值点作为候选匹配点, 得到匹配图像点集合为 $\{\mathbf{p}'_{l,i,k}\} (k \in [1, N_{l,i,k}], N_{l,i,k}$ 为采样点 $\mathbf{p}_{l,i}$ 对应的候选匹配点个数)。融合鲁棒估计的位姿估计问题可以描述为:

$$\hat{\boldsymbol{\mu}} = \underset{\boldsymbol{\mu}}{\operatorname{argmin}} \sum_{l=1}^{N_L} \sum_{i=1}^{N_{l,i}} \rho(d_{\mathbf{p}_{l,i}}(\mathbf{p}'_{l,i})^2),$$

$$d_{\mathbf{p}_{l,i}}(\mathbf{p}'_{l,i}) = \mathbf{n}_{l,i}^T (\mathbf{p}'_{l,i} - \Gamma(\mathbf{P}_{l,i}^M, \mathbf{E})), \quad (11)$$

其中: $\mathbf{p}'_{l,i}$ 为从候选匹配点集合 $\{\mathbf{p}'_{l,i,k}\}$ 中选取的最佳匹配点(选取策略为综合考虑模型直线投影的距离, 梯度值等), $\rho(\cdot)$ 为鲁棒估计算子, 最常用的是选择 RANSAC 算法进行鲁棒估计。

式(11)所描述的优化问题仅为每个模型直线采样点投影保留一个匹配点, 也就是每条模型直线仅有一条对应直线。当目标模型直线投影周围有多条目标强直线出现时, 如果选择的匹配直线对应错误, 有可能会引起后续跟踪的失败。为了提高跟踪的精度和鲁棒性, 有必要为每条模型直线保留多条候选直线(共线图像对应点集合), 并选取最优匹配直线作为模型直线对应。

对于每条投影直线 L_l , 通过 1D 搜索可以找到其上可见采样点对应的候选匹配点集合 $\{\{\mathbf{p}'_{l,i,k}\}\}$, 使用 RHT 可以在候选匹配点集合 $\{\{\mathbf{p}'_{l,i,k}\}\}$ 中拟合出 m_l 条直线, 即将 $\{\{\mathbf{p}'_{l,i,k}\}\}$ 分为 m_l 个共线点集合 $\{\mathbf{s}_1^l, \dots, \mathbf{s}_{m_l}^l\}$ 。由于模型直线的对应直线同模型直线方向近似平行, 因此在利用 RHT 进行聚类时, 限制同一个采样点 $\mathbf{p}_{l,i}$ 对应的候选匹配点 \mathbf{p}'_{l,i,k_1} 和 \mathbf{p}'_{l,i,k_2} 不能位于一个集合内

(即不共线)。利用 RHT 在剔除野值点的同时, 也将样本点分为多个共线点的集合, 即得到了每条模型直线对应的多条图像直线。

由于目标模型直线投影为图像上的直线, 考虑到图像畸变及其噪声的存在, 利用边缘点对直线进行拟合会存在拟合误差, 而拟合误差可以用来衡量直线的质量, 因而可以利用拟合误差对候选直线进行加权, 候选直线可以表示为 $\{(s_m^l, r_m^l)\}_{m=1}^{m_l}$, r_m^l 为集合 s_m^l 的拟合误差。每条候选直线的权重 $w_{m,1}^l$ 使用下式计算:

$$w_{m,1}^l = \begin{cases} e^{-\lambda_1 \left(\frac{r_m^l - r_{\min}^l}{r_{\max}^l - r_{\min}^l} \right)^2} & \text{if } r_{\max}^l \neq r_{\min}^l \\ 1 & \text{otherwise} \end{cases}, \quad (12)$$

其中 λ_1 为可设置参数, 根据需要进行设置。

另外最小二乘算法输入的初值为上一帧目标的位姿跟踪结果, 在相邻两帧内目标的位姿变换不会太剧烈, 因此模型直线投影同候选匹配直线的夹角和距离不会很大。综上引入以下权值:

$$w_{m,2}^l = \begin{cases} e^{-\lambda_2 \left(\frac{a_m^l - a_{\min}^l}{a_{\max}^l - a_{\min}^l} \right)^2} & \text{if } a_{\max}^l \neq a_{\min}^l \\ 1 & \text{otherwise} \end{cases}, \quad (13)$$

$$w_{m,3}^l = \begin{cases} e^{-\lambda_3 \left(\frac{d_m^l - d_{\min}^l}{d_{\max}^l - d_{\min}^l} \right)^2} & \text{if } d_{\max}^l \neq d_{\min}^l \\ 1 & \text{otherwise} \end{cases},$$

其中: a_m^l 为候选匹配直线同模型直线的夹角; d_m^l 为候选匹配直线同模型直线的端点距离最小值; λ_2, λ_3 为可设置参数。

综上候选匹配直线的权值可以设置为:

$$\mathbf{w}_m^l = w_{m,1}^l w_{m,2}^l w_{m,3}^l. \quad (14)$$

3.2 加权最小二乘算法

在实际情况中, 可能只有少数模型直线会出现一对多的情况, 可以采用遍历的方法(如果候选匹配直线组合数目很多, 则采用随机选取的方法), 利用式(15)计算每一种组合对应的目标位姿参数, 并选取具有最小重投影误差的匹配直线作为目标优化位姿:

$$\hat{\boldsymbol{\mu}} = \underset{\boldsymbol{\mu}}{\operatorname{argmin}} \sum_{l=1}^{N_L} \sum_{\substack{\mathbf{p}'_{l,i,k} \in \mathbf{s}_i^l}} \rho(d_{L_l}(\mathbf{p}'_{l,i,k})^2). \quad (15)$$

为了进一步抑制局部错误测量对结果的影响, 本文按照如下方法为每个采样点设置权值。

法向距离加权。边缘检测中的野值可能引起位姿估计值严重偏离真值, 为了降低野值点的影响, 权值设置为:

$$\omega_1 = \frac{1}{c + |d|}, \quad (16)$$

其中:常数 c 设置为所有法向距离的均方差, $|d|$ 为法向距离,该式的意义即采样点的权值与其法向距离近似成反比。

边缘强度加权。边缘梯度模值 g 可以作为一个有用的测度,因为弱的边缘很可能是不可靠的。本文用高低两个阈值来设置权值,即:

$$\omega_2 = \begin{cases} 0, & 0 \leq g < a \\ (g-a)/(b-a), & a \leq g < b \\ 1, & b \leq g \end{cases} \quad (17)$$

图像位置加权。在 3D 目标跟踪过程中,位于图像边界附近的采样点很有可能在下一帧图像中消失,这有可能会引起跟踪结果波动。因此可以根据采样点到图像边界的距离对其进行加权。

$$\omega_3 = \begin{cases} (h/2 - D)/D_{th}, & h/2 - D \leq D_{th} \\ 1, & h/2 - D > D_{th} \end{cases} \quad (18)$$

其中: h 为图像高度值; D 为采样点距离图像边界的最近距离; D_{th} 为距离阈值。由于图像的长宽比不一定为 1,为了保证上式对全图有效,本文对图像点进行坐标变换使得图像长宽比为 1。

综合以上各种考虑,设置每个采样点的权值为:

$$\omega = \omega_0 \omega_1 \omega_2 \omega_3, \quad (19)$$

其中 ω_0 是按照式(14)计算的采样点对应匹配点所在候选直线的权重。

因而通过迭代加权最小二乘(IRLS)最小化下式就可以实现参数求解:

$$\hat{\boldsymbol{\mu}} = \underset{\boldsymbol{\mu}}{\operatorname{argmin}} \sum_{i=1}^N (\omega_i \mathbf{n}_i^T (\mathbf{p}'_i - \Gamma(\mathbf{P}_i^M, \mathbf{E})))^2, \quad (20)$$

其解形式如下:

$$\hat{\boldsymbol{\mu}} = (\mathbf{J}^T \mathbf{W} \mathbf{J})^{-1} \mathbf{J}^T \mathbf{W} \mathbf{d}, \quad (21)$$

其中: \mathbf{W} 为 N 阶对角矩阵,对角线上第 i 个元素为第 i 个样本点的权重值。

3.3 算法流程

如图 3 所示,本文算法包括以下几个部分:

(a)图像预处理。根据像机内参和像差系数,对全图进行像差校正,校正后的图像将用于测量。

(b)模型更新。根据预测的目标位姿对模型直线进行消隐处理。

(c)模型采样点-图像点对应。对模型直线的投影进行等间隔采样,在图像中测量每个可见采

样点,并找到其对应的图像点,保留多个对应。

(d)多假设直线对应。对每条模型直线,使用 RHT 算法所拟合出的模型采样点对应多条直线,剔除野值点,并计算每条模型直线的权值。

(e)加权最小二乘优化。综合利用模型点的自身属性和所在直线的属性对模型采样点进行加权,并采用最小二乘方法优化模型点投影到图像点的距离,得到位姿优化结果。

(f)位姿参数解算。得到目标姿态矩阵后,通过分解旋转矩阵可以得到目标体的欧拉角,并得到目标体的姿态参数。

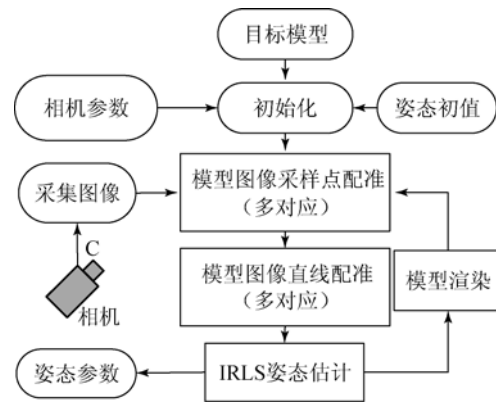


图 3 本文方法流程图

Fig. 3 Flowchart of proposed method

4 实验结果

本文主要研究的问题为航天器交会对接逼近段的相对姿态测量,因而实验中所使用的目标模型为比较通用的卫星模型(卫星主体加上两个太阳帆板)。本节实验分为三部分,首先利用数字仿真实验对本文算法进行精度验证,测试算法的有效性和估计的准确性;然后将本文基于多假设直线对应的优化算法同基于单假设对应的算法^[18]进行精度和鲁棒性的比较;最后设计实物仿真实验测试算法的鲁棒性。数字仿真实验采用 OpenGL 构建 3D 目标模型,仿真目标的外形如图 4 所示,目标体坐标系的原点位于立方体中心, x 轴同帆板长边平行(从左指向右), y 轴同帆板短边平行(从上指向下), z 轴满足右手系。从场景中获取的图像分辨率为 $640 \text{ pixel} \times 640 \text{ pixel}$,虚拟相机内参数为 ($C_x = C_y = 320, F_x = F_y = 1\ 814.8$),无像差。半实物仿真实验利用

相机拍摄目标模型,图像分辨率为 $640 \text{ pixel} \times 480 \text{ pixel}$,相机的内参数使用经典的棋盘格标定板进行标定^[20]。

4.1 精度分析实验

为了对本文算法进行精度验证,手工指定算法初值在真值附近加入较大的偏差,角度偏差为 $(-15^\circ, 15^\circ)$, x, y 方向的位置偏差为 $(-500 \text{ mm}, 500 \text{ mm})$, z 方向的位置偏差为 $(-2\ 000 \text{ mm}, 2\ 000 \text{ mm})$,然后利用算法的每一步输出结果将目标模型投影到图像上。由图 4,图 5(彩图见期刊电子版)可知,黄色直线为模型直线按照当前目标位姿在图像上的投影,青色细线为模型采样点到对应图像点的连线,青色粗线为模型采样点对应图像点的拟合直线。投影直线同图像目标直线越接近,青色细线越短,则解算误差越小。经过 8 次迭代后算法收敛。

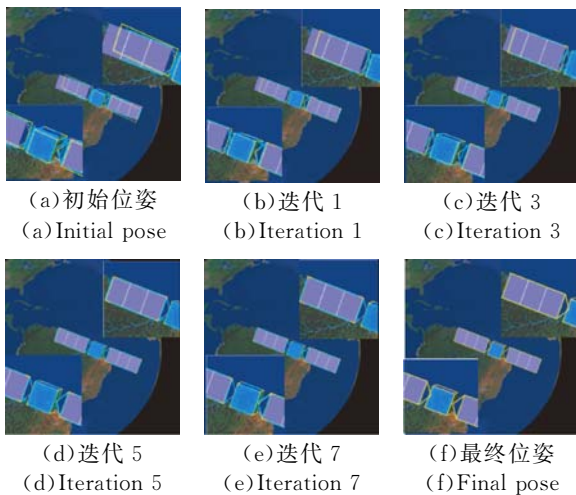


图 4 本文方法目标位姿估计结果 1

Fig. 4 Results of proposed method for frame 1

表 1 算法误差分析 1

Tab. 1 Pose error analysis of proposed method for frame 1

Pose parameters	Truth	Initial value	Optimized results	Error
$A_x/(\circ)$	-172.40	-164.07	-172.73	0.33
$A_y/(\circ)$	7.60	18.67	7.29	0.31
$A_z/(\circ)$	21.10	24.78	21.08	0.02
T_x/mm	341.61	379.39	330.50	11.11
T_y/mm	341.61	463.04	329.57	12.03
T_z/mm	46 078.38	47 888.56	46 113.51	35.13

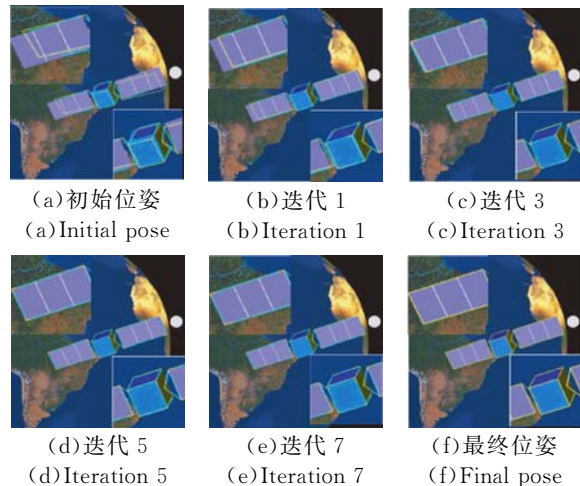


图 5 本文方法目标位姿估计结果 2

Fig. 5 Results of proposed method for frame 2

表 2 算法误差分析 2

Tab. 2 Pose error analysis of proposed method for frame 2

Pose parameters	Truth	Initial value	Optimized results	Error
$A_x/(\circ)$	-161.01	-152.68	-161.32	0.31
$A_y/(\circ)$	18.99	30.06	19.18	-0.19
$A_z/(\circ)$	17.88	14.19	-17.92	0.04
T_x/mm	400.35	438.14	387.18	3.17
T_y/mm	400.35	521.79	396.04	4.31
T_z/mm	33 016.60	34 826.78	32 986.11	30.49

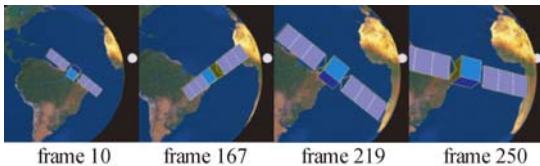
由表 1,表 2 可以看出,本文方法在垂直光轴方向的优化误差在 1 cm 左右,沿着光轴方向的优化误差为 3 cm 左右, x 方向角度估计误差为 0.5° 左右, y 方向角度估计误差为 0.3° 左右, z 方向估计精度优于 0.1° 。

4.2 精度对比实验

本小节利用 OpenGL 生成序列图像对两种方法的精度和鲁棒性进行综合比对。实验过程中,本文方法和单假设方法的参数保持不变,迭代次数都设置为 8 次,在每条模型直线上均等间隔采样 10 个模型采样点。目标沿光轴方向的运动距离为 $50 \sim 20 \text{ m}$,时间间隔为 0.1 s ,图像序列总时长为 30 s 。

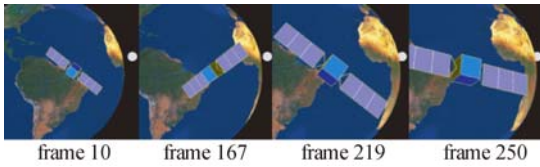
图 6(彩图见期刊电子版)给出了整个运动过程中基于单假设对应的方法跟踪出现一定偏差,而本文多假设方法跟踪成功的一些视图。图 6

中黄色轮廓线为目标模型在图像上的投影,位姿估计结果越准确,则其同图像上目标真实轮廓直线的位置越接近。由于本文方法在进行模型采样点和图像点匹配过程中,为每个模型采样点保留了多个假设,且将图像点约束在直线上,因而可以有效克服杂乱背景的干扰。



(a)单假设方法

(a)Single hypothesis based method



(b)本文方法

(b)Proposed multiple hypothesis based method

图 6 算法估计结果对比

Fig. 6 Comparison of pose estimation for simulation sequences

图 7 为两种算法的估计误差对比结果。

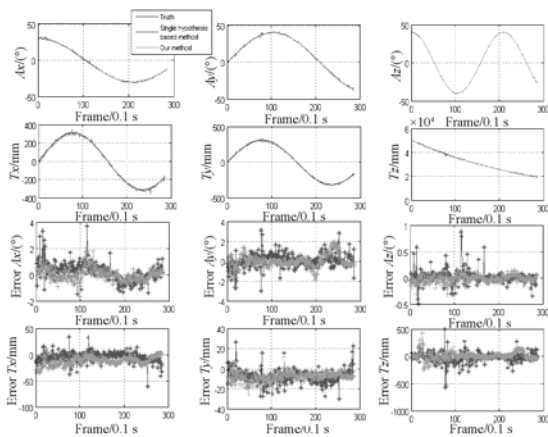


图 7 算法估计误差对比

Fig. 7 Pose estimation errors for simulation sequence

从图 7 可以看出,在整个运动过程中,基于单假设方法的运动轨迹上会出现多个毛刺,即位姿估计由于背景干扰而偏离真值;而本文方法估计的目标运动轨迹同目标真实运动轨迹的重合度

更好,轨迹更光滑,同时本文方法在每个时刻的偏差均略低于基于单假设方法(图 7 中本文的误差变化范围位于红色曲线之内)。

表 3 算法协方差误差分析

Tab. 3 Standard deviation of pose errors over simulation sequence

Error/ 1σ	$A_x/(\circ)$	$A_y/(\circ)$	$A_z/(\circ)$	T_x/mm	T_y/mm	T_z/mm
Single hypothesis based method	0.71	0.74	0.13	9.60	7.02	96.84
Our method	0.47	0.68	0.07	8.00	4.17	71.11

统计整个过程中两种方法的位姿估计值同真值的偏差可以得到表 3。由表 3 可以看出,本文方法在整个过程的跟踪精度相比基于单假设的方法有了一定提高。尽管本文方法跟踪精度提高幅度不是很大,但是本文方法有效抑制了杂乱背景的干扰,在整个过程中算法的跟踪稳定性显著提高。

本文方法的多假设实际上包括两部分:第一部分,每个模型采样点可能搜索到多个图像点对应(模型点配准阶段);第二部分,将搜索到的模型点对应聚类为多条直线,即为每条模型直线保留多条对应图像直线(模型直线配准阶段)。其中第一部分是第二部分的基础,只有在每个模型搜索到多个图像点对应的基础上,才能对共线的模型点对应进行聚类得到多条图像直线对应。本文方法可以看作局部和全局的结合,首先在模型点配准阶段为每个模型点保留多个可能的对应图像点(局部),然后在模型直线配准阶段将共线模型点对应的图像点集合聚类为多条直线,这一步也可看作是局部的(主要是为了剔除明显偏离直线的野值点),最后针对每一种可能组合解算位姿参数并选取误差最小的作为最终估计,这一步是全局的。因而本文算法相对于单假设方法受杂乱背景的干扰较小。

本文方法目前在台式机(i5 3.1GHz and 4GB RAM)进行了测试,对于文中 640×640 的图像(仿真目标含有 24 条直线段),1 s 可以处理 10 frame 左右,基本满足工程上的实时性需要。

4.3 实物验证实验

为了进一步验证方法的有效性,本小节在非

均匀强光照, 杂乱背景和部分遮挡条件下对算法进行了鲁棒性测试, 结果如图 8(彩图见期刊电子版)所示。本文方法对瞬间强光干扰(如图 8 第一行), 遮挡和杂乱背景(如图 8 第二行)以及旋转运动(如图 8 第三行)都具有很强的鲁棒性, 对 3D 直线进行的投影(图 8 中黄色轮廓线)表明, 本文方法可以鲁棒地跟踪目标的姿态。

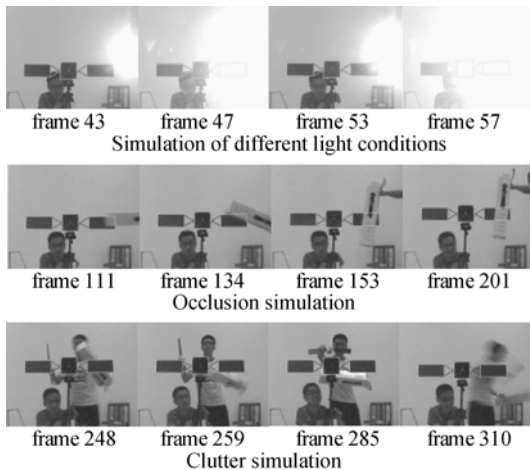


图 8 复杂环境条件下目标位姿估计结果

Fig. 8 Pose estimation for real sequence in complex environment

另外, 本节对本文算法的性能进行了定量衡量。实验中使用 4 张棋盘格图像进行相机内外参数标定(10×13 个内角点), 并求取所有棋盘格角点的重投影误差; 然后将棋盘格中的 23 条直线作为模型直线, 使用本文方法进行相机外参数解算, 计算重投影误差, 并同棋盘格方法^[20]进行比对。由于棋盘格角点可以获得亚像素级定位精度, 因此利用其计算得到的相机外参可以作为本文方法的参考值。由表 4 可以看出由本文方法得到的重投影误差同棋盘格方法相当, 因而本文方法可以获得同基于棋盘格方法相近的精度。

参考文献:

- [1] 耿明超, 赵铁石, 边辉, 等. 基于并联理论的单目视觉位姿估计[J]. 光学精密工程, 2013, 21(10): 2617-2626.
GENG M CH, ZHAO T SH, BIAN H, *et al.*. Pose estimation of monocular vision based on parallel theory [J]. *Opt. Precision Eng.*, 2013, 21(10): 2617-2626. (in Chinese)

表 4 棋盘格角点重投影误差比对

Tab. 4 Reprojection error comparison for chessboard corner

Reprojection Error/pixel	View 1	View 2	View3	View4
Corners based method	0.14	0.16	0.13	0.13
Proposed method	0.15	0.16	0.15	0.14

5 结 论

本文提出了一种基于多假设直线对应的目标位姿估计算法。为了有效克服杂乱背景或者特殊视图下模型-图像对应错误所引起的优化偏移问题, 本文算法在匹配阶段为每个模型采样点保留了多个图像点对应, 通过 RHT 算法将图像点对应聚类为多条直线, 并为每条模型直线保留多图像直线对应(多假设)。在样本点加权过程中, 算法综合利用样本点的局部特性和样本点的全局特性来提高算法对杂乱背景, 噪声以及特殊视角的鲁棒性。实验表明: 本文方法 X 方向角度估计误差优于 0.4°, y 方向优于 0.3°, z 方向优于 0.1°, 垂直光轴方向的相对位置误差优于 0.03%, 沿光轴方向的位置误差优于 0.1%。相比单假设方法, 本文方法能够有效克服复杂背景的干扰, 实现特殊视图目标位姿的稳定估计。

由于对每个模型采样点都要进行可见性判断, 1D 搜索以及雅克比矩阵计算, 因而当采样间隔较小, 模型采样点数目较大时, 算法的执行效率将降低。由于算法对每个模型点进行的操作可并行性较强, 同时目标颜色相对于背景具有很强的区分度, 因此未来工作中可以考虑结合直线和颜色模型进行目标位姿的估计, 同时利用硬件加速手段对算法进行并行加速。

- [2] 吕耀文, 王建立, 王昊京, 等. 应用抛物线运动估计摄像机姿态[J]. 光学精密工程, 2014, 22(4): 1078-1085
LV Y W, WANG J L, WANG H J, *et al.*. Estimation of camera poses by parabolic motion [J]. *Opt. Precision Eng.*, 2014, 22(4): 1078-1085. (in Chinese)
- [3] ZHANG L L, XU C, LEE K M, *et al.*. Robust and efficient pose estimation from line correspon-

- dences [C]. *Computer Vision-ACCV 2012 Lecture Notes in Computer Science*, 2013, 2013: 217-230.
- [4] MIRZAEI F M, ROUMELIOTIS S I. Globally optimal pose estimation from line correspondences [C]. *IEEE International Conference on Robotics and Automation (ICRA)*, 2011, 5581-5588.
- [5] ABABSA F, MALLEM M. Robust camera pose tracking for augmented reality using particle filtering framework [J]. *Machine Vision and Applications*, 2011, 22(1): 181-195.
- [6] KLEIN G, MURRAY D. Full-3D edge tracking with a particle filter [C]. *British Machine Vision Conference (BMVC)*, 2006, 1119-1128.
- [7] HARRIS C, STENNETT C. RAPID—a video rate object tracker [C]. *British Machine Vision Conference (BMVC)*, 1990, 73-78.
- [8] CHOI C, CHRISTENSEN H I. Real-time 3D model-based tracking using edge and keypoint features for robotic manipulation [C]. *2010 IEEE International Conference on Robotics and Automation (ICRA)*, 2010, 4048-4055.
- [9] VACCHETTI L, LEPETIT V, FUA P. Stable real-time 3D tracking using online and offline information [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2004, 26(10): 1385-1391.
- [10] PRESSIGOUT M, MARCHAND E. Real-time 3D model-based tracking: combining edge and texture information [C]. *IEEE International Conference on Robotics and Automation (ICRA)*, 2006, 2726-2731.
- [11] VACCHETTI L, LEPETIT V, FUA P. Combining edge and texture information for real-time accurate 3d camera tracking [C]. *Proceeding of the 3rd ISMAR'04 IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, 2004, 48-56.
- [12] PETIT A, MARCHAND E, KANANI K. A robust model-based tracker combining geometrical and color edge information [C]. *2913 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013, 3719-3724.
- [13] PANIN G, ROTH E, KNOLL A. Robust contour-based object tracking integrating color and edge likelihoods [C]. *Proceedings of the Vision, Modeling, and Visualization Conference*, 2008, 227-234.
- [14] MOUGHLBAY A A, CERVERA E, MARTINET P. Model based visual servoing tasks with an autonomous humanoid robot UR [J]. *Frontiers of Intelligent Autonomous Systems Studies in Computational Intelligence*, 2013, 466: 149-162.
- [15] EULIERE C, MARCHAND E, ECK L. Using multiple hypothesis in model-based tracking [C]. *IEEE International Conference on Robotics and Automation (ICRA)*, 2010, 4559-4565.
- [16] BROWN J A, CAPSON D W. A framework for 3D model-based visual tracking using a GPU-accelerated particle filter [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(1): 68-80.
- [17] PUPILLI M, CALWAY A. Real-time camera tracking using known 3d models and a particle filter [C]. *International Conference on Pattern Recognition (ICPR)*, 2006, 199-203.
- [18] DRUMMOND T, CIPOLLA R. Real-time visual tracking of complex structures [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2002, 24(7): 932-946.
- [19] 于起峰, 尚洋. 摄像测量学原理与应用研究[M]. 北京: 科学出版社, 2009.
- YU Q F, SHANG Y. *Videometrics: Principles and Researches* [M]. Beijing: Science Press, 2009. (in Chinese)
- [20] ZHANG Z Y. A flexible new technique for camera calibration [R]. *Microsoft Corporation: Technical Report, MSR-TR-98-71*, 1998.

作者简介:



张跃强(1987—),男,河南洛阳人,博士研究生,2009年于西安电子科技大学获得学士学位,2011年于国防科学技术大学获得硕士学位,主要从事机器视觉,摄像测量,图像分析等方面的研究。
E-mail: zyoungnudu@yahoo.com

导师简介:



于起峰(1958—),男,山东威海人,教授,中国科学院院士,博士生导师,1981年于西北工业大学获得学士学位,1983年于国防科学技术大学获得硕士学位,1995年于德国不莱梅大学获得博士学位,主要从事基于图像的精密测量与运动测量,光测实验力学与摄影测量,视觉导航与图像分析等方面的研究。
E-mail: yuqifeng@vip.163.com

(版权所有 未经许可 不得转载)