

## 基于 OPA 的时空约束单目视觉位姿估计算法

齐乃新<sup>a</sup>, 张胜修<sup>a</sup>, 曹立佳<sup>a</sup>, 杨小冈<sup>b</sup>, 陈世伟<sup>b</sup>  
(第二炮兵工程大学, a. 304 室; b. 303 室, 西安 710025)

**摘要:** 针对单目视觉系统中载体纯旋转运动引起的估计退化问题, 结合运动参数约束条件, 提出了一种基于 OPA 的时空约束单目视觉位姿估计方法。该方法利用 OPA 算法在估计本质矩阵和平移方向之前判断载体的运动形式, 检测修正估计退化问题, 并结合摄像机在运动过程中的时间相关约束及图像特征点在空间中的相关约束, 实现单目视觉系统的位姿估计。两组对比实验结果分析表明, 新算法相比于传统的估计算法具有更高的精度和稳定性。

**关键词:** 视觉导航; 单目视觉; OPA; 位姿估计; 时空约束

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1671-637X(2015)10-0039-05

## Temporal Constraint Pose Estimation Algorithm Based on OPA in Monocular Vision Systems

QI Nai-xin<sup>a</sup>, ZHANG Sheng-xiu<sup>a</sup>, CAO Li-jia<sup>a</sup>, YANG Xiao-gang<sup>b</sup>, CHEN Shi-wei<sup>b</sup>  
(The Second Artillery Engineering University, a. The 304 Room; b. The 303 Room, Xi'an 710025, China)

**Abstract:** Considering the estimation degradation caused by the pure rotational movement in the monocular vision system and the motion parameter constraints, we proposed an OPA-based spatial-time constraint pose estimation algorithm for monocular vision systems. The OPA algorithm was used to detect and correct the estimation degradation before estimating the essential matrix and translational direction vector. Combined with the time-related constraints of camera in motion process and the space-related constraints of image feature point, the estimation to navigation parameters of the monocular vision system were obtained. Comparative analysis of two sets of experimental results shows that the proposed algorithm has higher accuracy and stability compared with the conventional estimation algorithm.

**Key words:** visual navigation; monocular vision; OPA; pose estimation; temporal constraint

### 0 引言

视觉导航是根据摄像机采集的图像序列恢复摄像机的运动信息, 实现定位定姿的过程。位姿估计是指采用特定的算法求解不同参考系之间的相对位姿关系(包括3个旋转角和1个三维平移向量), 是视觉导航过程中的重要环节。就目前位姿估计方法的研究发展而言, 单目视觉算法较为成熟, 然而单目视觉位姿估计的过程是求解非线性问题, 实现起来较为复杂。因此, 寻求稳定高效的位姿估计方法仍然是国内外研究的热点问题。

早期的位姿估计算法<sup>[1]</sup>通过找到3个或4个不共

线的对应点, 采用最小二乘或直接矩阵求逆运算, 得到视觉系统唯一的位姿参数解。这种算法是根据世界坐标中的点和图像上的特征点的对应关系得到视觉系统的运动参数, 但位姿估计结果精度不高。随后, 位姿估计的方法产生了非线性迭代算法<sup>[2]</sup>和基于摄像机先验运动模型的滤波算法<sup>[3]</sup>两个方向。非线性迭代算法的典型代表是文献[4]中基于 Longuet-Higgins 准则的非线性优化方法, 具有较高的精度, 但针对非凸问题易陷入局部最优的困境。基于摄像机先验运动模型的滤波方法由于有先验运动模型不准的缺陷, 精度不高, 且可能带来维数灾难问题。

针对各类位姿估计方法的不足, 国内外学者做了大量的研究和改进, 取得了较多的研究成果。例如, 文献[5]针对深度相机, 提出了估计位姿变换精度的闭式解算法; 文献[6]提出了一种已标定和未标定摄像机同时适用的高精度快速算法, 用于解决对应点个数

收稿日期: 2014-06-27

修回日期: 2015-08-28

基金项目: 国家自然科学基金(61203189)

作者简介: 齐乃新(1989—), 男, 山东济南人, 硕士生, 研究方向为视觉导航。

大于等于6的摄像机位姿估计问题;文献[7]针对单目视觉系统的自运动估计问题,提出了滚动时域位姿估计算法。

上述改进算法在提高位姿估计的精度方面都取得了一定的成效,但是上述算法在估计载体的位姿参数时都没有考虑到运动过程中载体的纯旋转运动,忽视了图像序列间本质矩阵以及不同时刻摄像机平移方向的估计退化问题,影响了算法的计算效率和计算精度。针对这个问题,结合摄像机在运动过程中的时间相关约束及图像特征点在空间中的相关约束,提出了基于OPA的时空约束位姿估计算法,解决载体运动过程中的纯旋转运动估计退化问题。

## 1 摄像机模型及其运动估计算法

### 1.1 针孔成像模型

计算机视觉研究中,三维空间中的物体到像平面的投影关系即为成像模型,理想的投影成像模型是光学中的中心投影,也称为针孔模型,如图1所示。

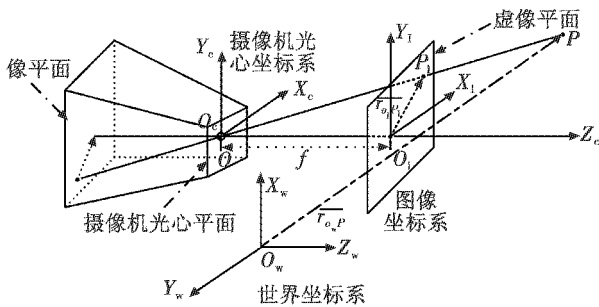


图1 针孔摄像机模型

Fig.1 Pinhole camera model

由光的直线传播条件可知,物体表面的反射光  $P$  经过一个针孔  $O$  投影到像平面上,对应地形成一个虚像平面,  $P_i$  为光线与虚像平面的交点。针孔  $O$  即为光心,过投影中心  $O$  且垂直于摄像机平面的射线称为光轴或主轴,主轴与虚像平面的交点称为摄像机的主点,光心与主点之间的距离定义为摄像机的焦距  $f$ 。由此可得到摄像机模型的描述为

$$z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R_{wc} & t_{wc} \\ \mathbf{0}_{1 \times 3} & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix}. \quad (1)$$

### 1.2 摄像机运动估计算法

现阶段,视觉导航大都是采用基于图像序列二维特征点匹配的位姿估计算法,总结起来有基于先验运动模型的滤波算法,基于两幅图像之间极线方程的本质矩阵特征值分解法<sup>[8]</sup>、最小二乘法以及非线性迭代算法<sup>[9]</sup>。两幅图像的极线几何关系如图2所示,对

应的极线方程可以写为

$$\tilde{m}_i^T (\mathbf{E} \cdot \tilde{m}_i) = 0 \quad (2)$$

式中:  $\mathbf{E} = \mathbf{t} \times \mathbf{R}$  是两幅图像对应的本质矩阵;  $\tilde{m}_i$  和  $\tilde{m}_i'$  分别是第一幅图像和第二幅图像上同一个三维特征点在空间中的归一化三维坐标。

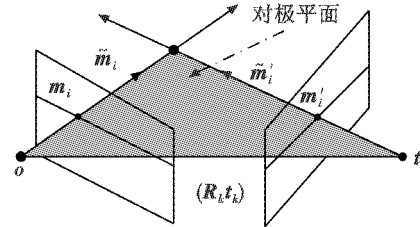


图2 极线几何关系示意图

Fig.2 The schematic of polar geometry

特征值分解方法和最小二乘法计算过程简单、实时性好,但是精度不高。非线性迭代方法是以式(2)为基础,通过构造式(3)所示的 Longuet-Higgins 准则直接估计出摄像机的未知运动参数,即

$$LH(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^n [\tilde{m}_i^T (\mathbf{t} \times \mathbf{R} \cdot \tilde{m}_i)]^2 \quad (3)$$

式中,匹配特征点个数用  $n$  表示。线性迭代方法能够很大程度上提高位姿估计的精度,但是容易陷入局部最优。基于摄像机运动模型的滤波估计方法能够解决最优解陷入局部最优问题,但是该方法位姿估计的精度不高,可能带来维数灾难的问题。

## 2 基于 OPA 的时空约束位姿估计算法

### 2.1 估计退化检测修正

载体的运动过程中会出现单纯平移和单纯旋转的情况,需要单独分析计算。在纯平移运动的情况下,第二幅图像中的点与第一幅图像中的点视场方向匹配,得到的3D点都移向(或远离)汇集点。此时所产生的本质矩阵  $\mathbf{E}$  在理想情况下是反对称的,可以直接估计出。在纯旋转运动的情况下,第二幅图像和第一幅图像之间不存在平移,导致本质矩阵  $\mathbf{E}$  和平移方向  $\mathbf{t}$  估计退化,最终导致位姿估计结果不准确,影响导航精度。同时,退化问题会使得目标函数在迭代过程中可能并不改进,一旦前面出现的基在迭代过程中又重新出现,则后面的迭代过程可能会在几个基上面兜圈子,出现“基的循环”现象,以致最终不能得到最优解。因此,在计算整个本质矩阵之前,检测载体的纯旋转运动,解决估计退化问题是十分必要的。

解决两幅图像之间的本质矩阵及平移方向估计退化问题的一个关键环节就是载体的纯旋转运动检测,当检测出载体的纯旋转运动后就可以通过纯旋转运动的本质矩阵求解方法来估计两幅图像之间的本质矩

阵,从而避免出现估计退化的问题。下文结合载体纯旋转运动的检测给出了估计退化检测修正的步骤。

1) 旋转矩阵估计。

旋转矩阵估计的一个常用方法是正交 Procrustes 算法(OPA)<sup>[10]</sup>,它计算 3×3 相关矩阵的 SVD 分解

$$C = \sum_i \tilde{m}'_i \tilde{m}_i^T = U \Sigma V^T \quad (4)$$

然后得到旋转矩阵  $R = UV^T$ 。

2) 旋转点  $\tilde{m}'_i$ 。

根据计算得到的旋转矩阵旋转点  $\tilde{m}'_i$ ,得到旋转后的点为

$$\hat{m}_i = R \tilde{m}'_i \quad (5)$$

3) 计算残差。

计算旋转后的点  $\hat{m}_i$  和点  $\tilde{m}'_i$  的残差

$$\Delta r = \sum_i \|\hat{m}_i - \tilde{m}'_i\|^2 \quad (6)$$

当得到的残差  $\Delta r$  大于某一个阈值时即可认为载体的运动不为纯旋转运动,可以进行下一步导航参数的估计。而当得到的残差  $\Delta r$  小于某一个阈值,即载体的运动为纯旋转运动的情况,此时并不存在平移,只需根据计算出的旋转矩阵计算导航参数,即可达到估计退化修正的目的。

2.2 时空约束估计算法

在载体不为纯旋转运动的情况下,要想估计出精度较高的导航参数,还需要寻求较好的估计算法。非线性迭代算法在具有完整约束条件的前提下能够很好地估计出载体的位姿参数,为此,本文在原有非线性迭代算法的基础上添加了导航系统在时间和空间上的约束条件。

1) 图像特征点空间相关约束。

极线约束是可以直接从两幅图像中恢复的,通常用式(2)的形式描述,然而,极线方程还具有其他的形式。在式(1)所表示的摄像机模型中,假设在第一个摄像机坐标系下三维点  $P_i$  的坐标为  $P_i = (x_i, y_i, z_i)^T$ ,在第二个摄像机坐标系下三维点  $P_i$  的坐标为  $P'_i = (x'_i, y'_i, z'_i)^T$ ,则有

$$\begin{cases} \tilde{m}'_i = P_i/z'_i \\ \tilde{m}_i = P'_i/z'_i \\ P_i = R P'_i + t \end{cases} \quad (7)$$

由式(7)可以得到

$$\tilde{m}'_i = (z_i R \cdot \tilde{m}_i + t)/z'_i \quad (8)$$

先用向量  $R \cdot \tilde{m}_i$  与式(8)叉乘,再用平移向量  $t$  和式(8)点乘,得到

$$t \cdot \tilde{m}'_i \times (R \cdot \tilde{m}_i) = 0 \quad (9)$$

式中,向量  $t, \tilde{m}'_i$  和  $R \cdot \tilde{m}_i$  在理想的情况下是共面的,式(2)和式(9)是代数等价的。

2) 时间相关约束。

不同时刻图像之间存在一定的时问相关约束,此约束即可作为约束条件用于摄像机运动参数的估计。

如图 3 所示,载体的运动过程是连续的,摄像机在不同时刻的位姿参数具有密切的联系关系而非独立的,其思想类似于环闭检测,称为虚拟闭环,不需要摄像机进行真实环闭运动,累积误差通过不相邻时刻场景特征的相似性进行消减。

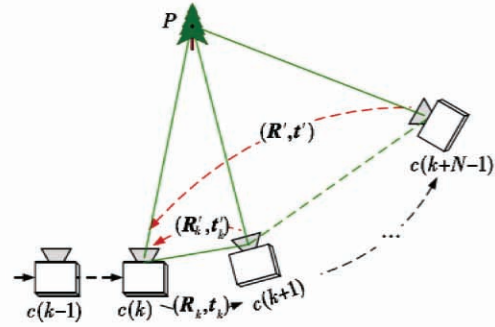


图 3 时间相关位姿约束关系示意图

Fig. 3 Sketch map for time corrected constraints

图中,  $R^T = R_{c(k)c(k+N-1)}$ ,  $t^T = t_{c(k)c(k+N-1)c(k)}$ 。首先,假设任意相邻两幅图像的匹配点对为  $\tilde{m}_i, \tilde{m}'_i, i = 1, \dots, n_k$ 。该优化函数存在如式(10)所示的约束关系。为了方便叙述,令  $R_k = R_{c(k)c(k+1)}$ ,  $R_k^T = R_{c(k+1)c(k)}$ ,  $t_k = t_{c(k+1)c(k)c(k+1)}$ ,  $t_k^T = t_{c(k)c(k+1)c(k)}$ , 则有

$$\begin{cases} t_k \cdot \tilde{m}'_i \times (R_k \cdot \tilde{m}_i) = 0 \\ t_k^T \cdot \tilde{m}_i \times (R_k^T \cdot \tilde{m}'_i) = 0 \\ R_k^T \cdot R_k = I(3) \\ R_k \cdot t_k^T = -t_k \end{cases} \quad (10)$$

假设在载体的整个运动过程中摄像机总共采集到了  $T$  帧图像,则根据摄像机成像模型和极约束条件可以得到所有  $T$  帧图像之间约束条件。如图 3 所示,图像间的旋转矩阵满足

$$R_{c(1)c(T)} = R_{c(T-1)c(T)} \cdot R_{c(T-2)c(T-1)} \cdot \dots \cdot R_{c(1)c(2)} \quad (11)$$

而摄像机平移向量关系为

$$t_{c(1)c(T)w} = t_{c(2)c(1)w} + t_{c(3)c(2)w} + \dots + t_{c(T)c(T-1)w} \quad (12)$$

由以上约束条件可以得到摄像机在整个运动过程中,用于位姿参数最优估计的非线性优化函数为

$$(R_1, t_1, \dots, R_T, t_T) = \arg \left\{ \min \left[ \sum_{j=1}^T \left[ \sum_{i=1}^{n_j} (\tilde{m}'_{ij} \cdot (t_j \times R_j \cdot \tilde{m}_{ij})) \right] \right] \right\} \quad (13)$$

$$s. t. \quad \prod_{j=1}^T R_j = R_{c(1)c(T)},$$

$$\sum_{j=1}^T (\mathbf{R}_{c(j)w} \cdot \mathbf{t}_j) = \mathbf{R}_{c(1)w} \cdot \mathbf{R}_{c(1)c(T)c(1)} \circ$$

在载体的整个运动过程中,摄像机位姿的最优估计结果满足式(10)和式(13)建立的约束条件。将此约束条件与传统的极线约束条件、图像特征点空间相关约束条件综合用于非线性迭代的优化目标函数,能够精确地估计载体的运动参数。

### 3 实验及结果分析

实验设计了原地旋转实验和行驶导航实验,其中原地旋转实验专门用于评估载体姿态测量的精度,行驶导航实验通过与其他几种位姿估计算法相比,综合评估载体在实际运动中的姿态测量精度。

本实验中所用的实验平台是根据需要自主改装的由DrRobot公司生产的X80-H型WiFi控制智能小车,如图4所示。该智能实验平台包含视觉传感器和通信控制模块,能够通过编程预定行驶轨迹模拟平台的运动真值,并且能够通过WiFi实时传输图像序列,用于估计平台的位姿参数。实验选在实验室环境,摄像机的视场中有各种实验器材、办公用品等静态的景物特征,并且还含有人员走动等动态元素。在此复杂的环境下,本文进行了原地旋转和L形行走转弯两组实验。

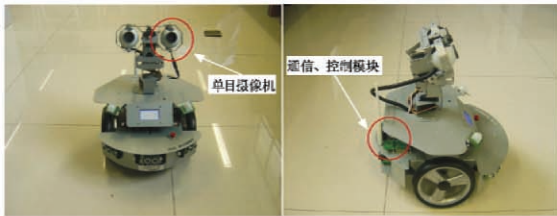


图4 智能实验平台

Fig. 4 Smart experimental platform

#### 实验1 原地旋转实验。

在原地旋转实验中,单目视觉系统连续转动90°,摄像机以20帧/s的频率采集图像序列,分别用本质矩阵特征值分解法、非线性迭代算法以及基于OPA的时空约束导航算法估计平台的位姿参数,进行对比评估。实验对应的位姿计算结果如图5所示。

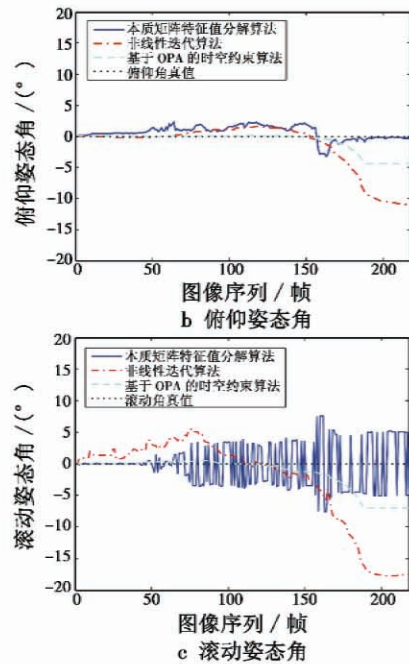
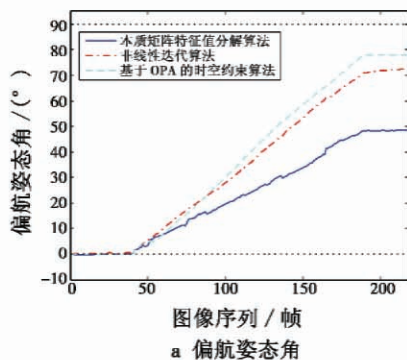


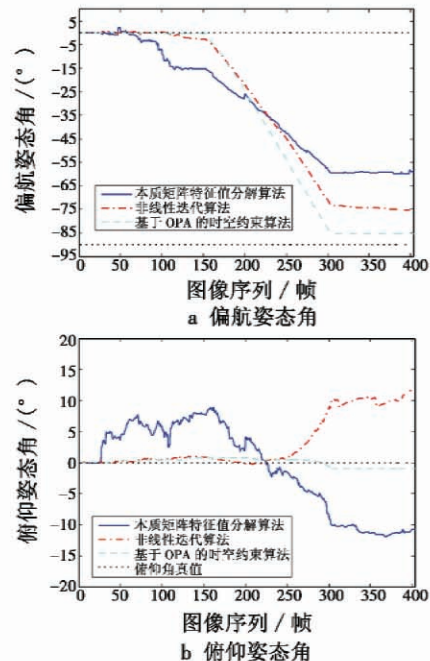
图5 实验1位姿计算结果

Fig. 5 Pose calculation result in Experiment 1

从图5可以看出,在原地旋转90°的实验条件下,3种导航位姿估计算法都能估计出载体的姿态角信息,但是相比于传统的本质矩阵特征值分解算法和非线性迭代算法,本文提出的基于OPA的时空约束导航算法具有明显的优势,姿态测量精度更高。

#### 实验2 L形行走转弯实验。

L形行走转弯实验的实验过程由东向西行驶,90°转弯,向南行驶3部分组成。实验中分别采用本质矩阵特征值分解法、非线性迭代算法以及基于OPA的时空约束导航算法进行实验分析,实验结果如图6所示。



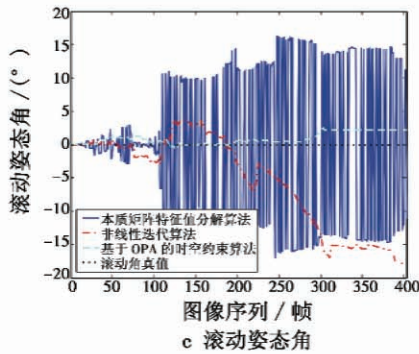


图 6 实验 2 位姿计算结果

Fig. 6 Pose calculation result in Experiment 2

从图 6 可以看出,在未知环境导航的实验条件下,相比于传统的视觉导航位姿估计算法,基于 OPA 的时空约束导航算法具有更高的精度,偏航姿态角误差能够控制在  $5^\circ$  以内,俯仰姿态角误差能够控制在  $3^\circ$  以内,滚动姿态角误差能够控制在  $4^\circ$  以内。

通过两组对比实验,可以得出基于 OPA 的时空约束导航算法在纯旋转运动时能够解决估计退化的问题,较大程度上提高了姿态估计的精度,能够较好地应用于未知环境下载体的姿态估计。

#### 4 结论

针对单目视觉导航算法中的位姿估计问题,介绍了摄像机的成像模型,在描述单目视觉导航基本原理和具体过程的前提下,采用 OPA 算法解决了估计退化问题,通过建立摄像机运动的时间域和空间域的约束条件,提出了基于 OPA 的时空约束单目视觉位姿估计算法。最后通过两组实验,验证了提出的视觉导航算法相比于传统的算法具有更高的精度和稳定性。

#### 参考文献

[1] DHOME M, RICHTIN M. Determination of the attitude of 3D objects from a single perspective view[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(12):1265-1278.

[2] KLEIN G, MURRAY D. Improving the agility of key frame

based SLAM[C]//Proceedings of the 10th European Conference on Computer Vision, Mar-seille, Germany, 2008: 802-815.

[3] DAVISON A, MOLTON N, REID I. Mono SLAM: real-time single camera SLAM[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6):1052-1067.

[4] LONGUET H C, HIGGINS F R S. Multiple interpretations of a pair of images of a surface[C]//Proceedings of the Royal Society of London Series: A Mathematical Physical and Engineering Sciences, 1988:1-15.

[5] 李兴东,郭伟,李满天,等.一种估计深度相机位姿精度的闭式算法[J].机器人,2014(2):194-202,209. (LI X D, GUO W, LI M T, et al. A closed-form solution for estimating the accuracy of depth camera's relative pose[J]. Robot, 2014(2):194-202, 209.)

[6] 李书杰,刘晓平.摄像机位姿的高精度快速求解[J].中国图象图形学报,2014(1):20-27. (LI S J, LIU X P. An accurate and fast algorithm for camera pose estimation[J]. Journal of Image and Graphics, 2014(1):20-27.)

[7] 杨东方,孙富春,王仕成.单目视觉系统自运动的滚动时域估计算法[J].模式识别与人工智能,2013(1):76-82. (YANG D F, SUN F C, WANG S C. Moving horizon estimation of ego-motion in monocular visual systems[J]. Pattern Recognition and Artificial Intelligence, 2013(1):76-82.)

[8] MCINROY J, QI Z. A novel pose estimation algorithm based on points to regions correspondence[J]. International Journal of Math Imaging, 2008, 30(2):195-207.

[9] HIGGINS H L. Multiple interpretations of a pair of images of a surface[C]//Proceedings of the Royal Society of London Series: A Mathematical Physical and Engineering Sciences, 1988:1-15.

[10] 周拥军,寇新建.正交 Procrustes 分析及其在旋转矩阵估计中的应用[J].武汉大学学报:信息科学版,2009(8):996-999. (ZHOU Y J, KOU X J. Orthogonal Procrustes analysis and its application on rotation matrix estimation[J]. Geomatics and Information Science of Wuhan University, 2009(8):996-999.)



请扫描二维码关注我刊