文章编号:1002-2082(2019)01-0253-06

基于机器视觉的最大似然位姿估计算法

屈也频,张超然,吕余海

(海军研究院,上海 200235)

摘 要:针对现有位姿估计算法对采样数据不做任何的统计假设,缺少评判标准等问题,从信号的概率密度函数出发,推导了基于机器视觉的最大似然位姿估计的一般形式,并证明利用单幅图像时,在各向同性高斯噪声情况下传统迭代算法与最大似然估计等效。推导了位姿估计的克拉美-罗界,给出了位姿估计的方差下限。根据仿真结果可以看出,利用 10 张图像时,最大似然算法在噪声功率大于 5 dB 的情况下,性能明显优于传统迭代算法,证明适当增加图像数可有效提高估计性能。

关键词:机器视觉;位姿估计;最大似然估计;克拉美-罗界

中图分类号:TN201; TP391.41 **文献标志码**:A **DOI**:10.5768/JAO201940.0202002

Maximum likelihood pose estimation using machine vision

QU Yepin, ZHANG Chaoran, LYU Yuhai

(Naval Academy, Shanghai 200235, China)

Abstract: The existing pose estimation algorithms do not make any statistical assumptions on the sampled data, and lack the evaluation criteria. Aiming at this problem, based on the probability density function of the signal, we derived the general form of maximum likelihood pose estimation based on machine vision was and proved that the traditional iterative algorithm is equivalent to the maximum likelihood estimation using single imag ein the case of isotropic Gaussian noise. What's more, we derived the Cramér-Rao bound of pose estimation, which could be regarded as the variance low bound of any unbiased estimations. By the analysis of the simulation, the maximum likelihood method is much better than the traditional iterative method by using 10 pictures when noise power is greater than 5 dB, it proves that the performance of pose estimation can be improved by increasing the number of images.

Key words: machine vision; pose estimation; maximum likelihood estimation; Cramér-Rao bound

引言

利用相机对位置已知的控制点进行拍摄,通过拍摄得到的二维图像解算相机在世界坐标系下的位置和姿态,称为位姿估计问题,是机器人导航、计算机视觉等领域[1-3]的核心问题之一。位姿估计根据计算方法分为 2 种方式[4]:第一种一般通过解算一组与旋转参数和平移参数相关的多项式,进而获得位姿估计结果[5],这类方法对于噪声敏感[3],估计性能有限;另一种通常称为迭代类算法,这种算法通过建立代价函数[6],将位姿估计问题转换为非线性最小二乘优化问题,然后利用 Gauss-Newton[7]、Levenberg-Marquardt 等方法进行位姿参数的估计,这类算法能够综合利用多点的冗余信息[8],增加了估计的鲁棒性,迭代算法中还有一种称为正交迭代(orthogonal iteration, OI)算法,该算法将物空间共线性误差的最小化视为

收稿日期:2018-07-24; **修回日期**:2018-09-10

基金项目:国防预研基金项目(4010905010301)

作者简介:屈也频(1962—),男,博士,研究员,博士生导师,主要从事航空装备研究。E-mail:qypin@126.com

求解绝对定向问题，然后利用迭代方式进行参数估计[9]，这种方法计算量较小，速度较快[10]，并保证了全局收敛性，但是性能略逊于传统迭代算法。

传统迭代算法是从信号的测量误差出发[11]，利用误差最小化的原则建立代价函数，从而得到估计值。这一类算法是一种最小二乘估计，试图使采样得到的信号和无噪声情况下的数据之差的平方达到最小，这一类方法对采样数据不做任何的统计假设，性能取决于噪声的特性，并且往往不是最佳估计，而由于没有对信号做任何统计假设，估计的统计性能是无法评价的[12]。由于相机拍摄的不确定性，图片采样所得的像素信号应视为随机信号，本文从随机信号的统计特性出发，根据采样信号的概率密度推导了最大似然位姿估计的一般形式。从理论上证明了利用单幅图像，在各向同性高斯噪声情况下传统迭代算法等效于最大似然估计，因此在理论上传统迭代算法只能应用于各向同性高斯噪声，对于复杂噪声，将会产生模型误差，导致算法失效。本文的最大似然算法将是推导复杂噪声下的位姿估计的基础；另外，根据信号的 Fisher 信息阵推导了位姿估计的克拉美-罗界(Cramér-Rao bound, CRB)，作为任何无偏估计的方差下界，克拉美-罗界可以作为位姿估计方法有效性的评价标准，并且根据所得结果可知，通过适当增加采样个数可以有效提高估计的性能。这是由于通过采样信号个数的增加，可以更好地估计随机噪声的统计特性，从而减小随机噪声对于估计的影响。

1 信号模型

假设有 n 个控制点，其在世界坐标系中的坐标为 p_i^w，$i=1,2,\cdots,n$，它们在相机坐标系下的坐标为 p_i^c，$i=1,2,\cdots,n$，相机对控制点进行拍摄，可以得到像平面上的 n 个像点，像点在图像坐标系下的物理坐标为 p_i^p，$i=1,2,\cdots,n$，在图像坐标系下的像素坐标为$(u_i,v_i)^{\mathrm{T}}$，如图 1 所示。在 k 时刻对控制点进行拍摄，考虑采样过程的随机噪声影响，则观测量为如下随机向量：

$$s(k)=[\hat{u}_1(k),\hat{v}_1(k),\hat{u}_2(k),\hat{v}_2(k),\cdots,\hat{u}_n(k),\hat{v}_n(k)]^{\mathrm{T}}=\boldsymbol{m}(k)+\boldsymbol{n}(k) \tag{1}$$

式中：$k=1,2,\cdots,K$，其中：

$$\boldsymbol{m}(k)=[u_1(k),v_1(k),u_2(k),v_2(k),\cdots,u_n(k),v_n(k)]^{\mathrm{T}} \tag{2}$$

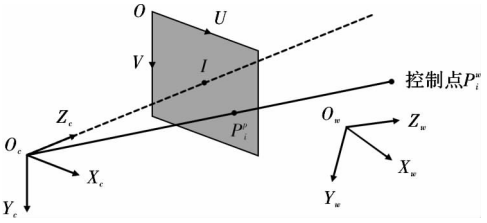


图 1 位姿估计问题的几何示意图

Fig. 1 Geometric sketch of pose estimation

世界坐标与相机坐标存在以下关系：

$$\boldsymbol{p}_i^c(k)=\boldsymbol{R}(k)\boldsymbol{p}_i^w(k)+\boldsymbol{t}(k) \tag{3}$$

由共线方程可以得到：

$$\begin{cases}\dfrac{x_i^p(k)-x_o^p}{f}=\dfrac{x_i^c(k)}{z_i^c(k)}\\[2ex]\dfrac{y_i^p(k)-y_o^p}{f}=\dfrac{y_i^c(k)}{z_i^c(k)}\end{cases} \tag{4}$$

式中$(x_o^p,y_o^p)^{\mathrm{T}}$为图像主点 I 在图像坐标系下的物理坐标，则像素坐标系与世界坐标系的转换关系可以表示为

$$\begin{cases}u_i(k)=\dfrac{f}{\mathrm{d}_x}\dfrac{\boldsymbol{r}_1(k)\boldsymbol{p}_i^w(k)+t_x(k)}{\boldsymbol{r}_3(k)\boldsymbol{p}_i^w(k)+t_z(k)}+\dfrac{x_o^p}{\mathrm{d}_x}\\[2ex]v_i(k)=\dfrac{f}{\mathrm{d}_y}\dfrac{\boldsymbol{r}_2(k)\boldsymbol{p}_i^w(k)+t_y(k)}{\boldsymbol{r}_3(k)\boldsymbol{p}_i^w(k)+t_z(k)}+\dfrac{y_o^p}{\mathrm{d}_y}\end{cases} \tag{5}$$

式中：$t_x(k)$、$t_y(k)$、$t_z(k)$分别为 $\boldsymbol{t}(k)$从上到下的 3 个元素；$\boldsymbol{r}_1(k)$、$\boldsymbol{r}_2(k)$、$\boldsymbol{r}_3(k)$分别为 $\boldsymbol{R}(k)$的从上到下的 3 个横向量；$\boldsymbol{R}$ 可以用 3 个欧拉角表示。假设在 K 次拍摄过程中，相机的姿态、控制点的世界坐标保持不变，即 $\boldsymbol{m}$ 不随时间变化，则采样信号可以表示为

$$\boldsymbol{s}(k)=\boldsymbol{m}+\boldsymbol{n}(k) \tag{6}$$

2 最大似然估计

最大似然估计是一种使用了信号概率模型的方法，其基本思想是参数的估计值应是使观测信号概率最大的值。假设信号为高斯随机变量，噪声的均值为零，则 $\boldsymbol{m}$ 即为观测量 $\boldsymbol{s}(k)$的均值，根据概率论可知，k 时刻观测量的概率密度为[13]

$$p_{s(k)}=\frac{1}{\det(\pi\boldsymbol{Q}_s)}\exp\{-[\boldsymbol{s}^H(k)-\boldsymbol{m}^H]\boldsymbol{Q}_s^{-1}[\boldsymbol{s}(k)-\boldsymbol{m}]\} \tag{7}$$

式中：$\det(\cdot)$表示矩阵的行列式；$\boldsymbol{Q}_s$ 为随机信号 $\boldsymbol{s}$ 的方差。与待估计参数无关，待估计参数表示为

$$\boldsymbol{\theta}=[\alpha,\beta,\gamma,t_x,t_y,t_z]^{\mathrm{T}} \tag{8}$$

则 $\boldsymbol{m}(\boldsymbol{\theta})$表示为 $\boldsymbol{\theta}$ 的函数，假设 K 次拍摄是统计独立的，则 K 次拍摄观测值的联合密度函数为

$$p_{s(1),s(2),\cdots,s(K)}=\prod_{k=1}^{K}\frac{1}{\det(\pi\boldsymbol{Q}_s)}\exp\{-[\boldsymbol{s}^H(k)-\boldsymbol{m}^H(\boldsymbol{\theta})]\boldsymbol{Q}_s^{-1}[\boldsymbol{s}(k)-\boldsymbol{m}(\boldsymbol{\theta})]\} \tag{9}$$

对等号两边求自然对数，去掉常数项并除以 K，得到似然函数为

$$L(\boldsymbol{\theta})=-\ln[\det(\boldsymbol{Q}_s)]-\frac{1}{K}\sum_{k=1}^{K}[\boldsymbol{s}^H(k)-\boldsymbol{m}^H(\boldsymbol{\theta})]\boldsymbol{Q}_s^{-1}[\boldsymbol{s}(k)-\boldsymbol{m}(\boldsymbol{\theta})] \tag{10}$$

(10)式的第二项可以表示为

$$\frac{1}{K}\sum_{k=1}^{K}[\boldsymbol{s}^H(k)-\boldsymbol{m}^H(\theta)]\boldsymbol{Q}_s^{-1}[\boldsymbol{s}(k)-\boldsymbol{m}(\boldsymbol{\theta})]=$$
$$\mathrm{tr}\{\frac{1}{K}\sum_{k=1}^{K}[\boldsymbol{s}^H(k)-\boldsymbol{m}^H(\theta)]\boldsymbol{Q}_s^{-1}[\boldsymbol{s}(k)-\boldsymbol{m}(\boldsymbol{\theta})]\}=$$
$$\mathrm{tr}\{\boldsymbol{Q}_s^{-1}\frac{1}{K}\sum_{k=1}^{K}[\boldsymbol{s}(k)-\boldsymbol{m}(\boldsymbol{\theta})][\boldsymbol{s}^H(k)-\boldsymbol{m}^H(\boldsymbol{\theta})]\}=$$
$$\mathrm{tr}\{\boldsymbol{Q}_s^{-1}\boldsymbol{C}_s(\boldsymbol{\theta})\} \tag{11}$$

式中：tr{·}表示取矩阵的迹；$\boldsymbol{C}_s(\boldsymbol{\theta})$定义为信号相关矩阵：

$$\boldsymbol{C}_s(\theta)\overset{\Delta}{=}\frac{1}{K}\sum_{k=1}^{K}[\boldsymbol{s}(k)-\boldsymbol{m}(\boldsymbol{\theta})][\boldsymbol{s}^H(k)-\boldsymbol{m}^H(\boldsymbol{\theta})] \tag{12}$$

于是得到似然函数：

$$L(\theta)=-\ln[\det(\boldsymbol{Q}_s)]-\mathrm{tr}\{\boldsymbol{Q}_s^{-1}\boldsymbol{C}_s(\boldsymbol{\theta})\} \tag{13}$$

(13)式为最大似然估计的一般形式，使似然函数最大的参数 θ 即为位姿估计的最大似然估计，可以通过搜索或者更高效的迭代算法计算。

当噪声为各向同性的高斯噪声时，$\boldsymbol{Q}_s=\delta_w^2\boldsymbol{I}$，$\delta_w^2$为噪声功率，似然函数可以表示为

$$L(\theta)=-2n\ln\delta_w^2-\frac{1}{\delta_w^2}\mathrm{tr}\{C_s(\boldsymbol{\theta})\} \tag{14}$$

由于似然函数取最大值的必要非充分条件为

$$\frac{\partial}{\partial\delta_w^2}L(\boldsymbol{\theta})=0 \tag{15}$$

得到：

$$\delta_w^2=\frac{\mathrm{tr}\{\boldsymbol{C}_s(\boldsymbol{\theta})\}}{2n} \tag{16}$$

将上式代入似然函数，去除常数项并除以 $2n$，似然函数化简为

$$L(\boldsymbol{\theta})=-\ln\mathrm{tr}\{\boldsymbol{C}_s(\boldsymbol{\theta})\} \tag{17}$$

根据 $\boldsymbol{C}_s(\boldsymbol{\theta})$的定义，求似然函数的最大值等效于求下式的最小值：

$$\mathrm{tr}\{\boldsymbol{C}_s(\boldsymbol{\theta})\}=\frac{1}{K}\sum_{k=1}^{K}[\boldsymbol{s}^H(k)-\boldsymbol{m}^H(\boldsymbol{\theta})][\boldsymbol{s}(k)-\boldsymbol{m}(\boldsymbol{\theta})]=$$
$$\frac{1}{K}\sum_{k=1}^{K}\sum_{1i=1}^{n}\left[\left(u_i(k)-\frac{f}{d_x}\frac{\boldsymbol{r}_1(k)\boldsymbol{p}_i^w(k)+t_x(k)}{\boldsymbol{r}_3(k)\boldsymbol{p}_i^w(k)+t_z(k)}-\frac{x_o^p}{d_x}\right)^2+\left(v_i(k)-\frac{f}{d_y}\frac{\boldsymbol{r}_2(k)\boldsymbol{p}_i^w(k)+t_y(k)}{\boldsymbol{r}_3(k)\boldsymbol{p}_i^w(k)+t_z(k)}-\frac{x_o^p}{d_y}\right)^2\right] \tag{18}$$

利用单幅图像时，上式与传统的迭代算法等效，因此，利用单幅图像时，在各向同性高斯噪声情况下传统迭代算法与最大似然算法等效。值得注意的是，上述结果是在(16)式的情况下得到的，即位姿估计是在噪声功率估计的基础上实现的，所以增加信号的采样，提高对噪声统计特性的估计，有利于提高参数估计的性能。

3 克拉美-罗界分析

为了了解参数估计的潜在性能，可以利用克拉美-罗界作为评价标准。克拉美-罗界给出了参数所有无偏估计的方差极限[14]，将估计误差的方差表示为 $\boldsymbol{C}(\boldsymbol{\theta})$，对于所有无偏估计，有关系式：

$$\boldsymbol{C}(\boldsymbol{\theta})\geqslant\boldsymbol{C}_{\mathrm{CR}}(\boldsymbol{\theta}) \tag{19}$$

式中 $\boldsymbol{C}_{\mathrm{CR}}(\boldsymbol{\theta})$表示克拉美-罗界，克拉美-罗界可以通过求解 Fisher 信息阵 $\boldsymbol{J}$ 的逆矩阵求得：

$$\boldsymbol{C}_{CR}(\boldsymbol{\theta})=\boldsymbol{J}^{-1} \tag{20}$$

Fisher 信息阵 $\boldsymbol{J}$ 中的元素为

$$[\boldsymbol{J}]_{i,j}=E\left[\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}\right] \tag{21}$$

通过对(13)式求偏导并取期望可以得到$[\boldsymbol{J}]_{i,j}$通用的形式(推导可参见文献[15])：

$$[\boldsymbol{J}]_{i,j}=K\mathrm{tr}\left[\boldsymbol{Q}_s^{-1}\frac{\partial\boldsymbol{Q}_s}{\partial\theta_i}\boldsymbol{Q}_s^{-1}\frac{\partial\boldsymbol{Q}_s}{\partial\theta_j}\right]+2K\mathrm{Re}\left[\frac{\partial\boldsymbol{m}^H(\theta)}{\partial\theta_i}\boldsymbol{Q}_s^{-1}\frac{\partial\boldsymbol{m}(\theta)}{\partial\theta_j}\right] \tag{22}$$

(22)式为 Fisher 信息阵的一般化形式，是许多克拉美－罗界推导的起点，需要注意的是，通常方差 $\boldsymbol{Q}_s$ 是一个未知矩阵，因此上式应该包含方差的未知参数。

在位姿估计问题中，由于(22)式的第 1 项只与方差所含未知参数有关，第 2 项只与待估计参数有关，可以将 Fisher 信息阵分块：

$$\boldsymbol{J}=\begin{bmatrix}\boldsymbol{J}_{\boldsymbol{Q}_s\boldsymbol{Q}_s} & 0\\ 0 & \boldsymbol{J}_{00}\end{bmatrix} \tag{23}$$

$$[\boldsymbol{J}_{\boldsymbol{Q}_s\boldsymbol{Q}_s}]_{i,j}=K\mathrm{tr}\left[\boldsymbol{Q}_s^{-1}\frac{\partial\boldsymbol{Q}_s}{\partial q_i}\boldsymbol{Q}_s^{-1}\frac{\partial Q_s}{\partial q_j}\right] \tag{24}$$

$$[\boldsymbol{J}_{00}]_{i,j}=2K\mathrm{Re}\left[\frac{\partial\boldsymbol{m}^H(\boldsymbol{\theta})}{\partial\theta_i}Q_s^{-1}\frac{\partial\boldsymbol{m}(\boldsymbol{\theta})}{\partial\theta_j}\right] \tag{25}$$

式中 q_i 为方差所含的未知参数，当噪声为各点无关的各向同性高斯噪声时，$\boldsymbol{Q}_s=\delta_w^2\boldsymbol{I}$，方差所含未知参数为 δ_w^2，则有：

$$\boldsymbol{J}_{Q_sQ_s}=\boldsymbol{J}_{\delta_w^2\delta_w^2}=\frac{2nK}{(\delta_w^2)^2} \tag{26}$$

根据矩阵分块求逆原理，噪声功率估计的克拉美-罗界为

$$\boldsymbol{C}_{CR}(\delta_w^2)=\frac{(\delta_w^2)^2}{2nK} \tag{27}$$

可以看出噪声功率估计的克拉美-罗界随着采样数的增加而减小，说明增加信号的采样，可以提高对噪声统计特性的估计。将 $\boldsymbol{Q}_s=\delta_w^2\boldsymbol{I}$ 代入(25)式，得到：

$$[J_{\theta\theta}]_{i,j}=\frac{2K}{\delta_w^2}\mathrm{Re}\left[\frac{\partial \boldsymbol{m}^{\mathrm{H}}(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial \boldsymbol{m}(\boldsymbol{\theta})}{\partial\theta_j}\right] \tag{28}$$

定义

$$D\xlongequal{\Delta}\frac{\partial \boldsymbol{m}^{\mathrm{H}}(\boldsymbol{\theta})}{\partial\theta} \tag{29}$$

由于 $\boldsymbol{\theta}=[\alpha,\beta,\gamma,t_x,t_y,t_z]^{\mathrm{T}}$，待估计参数的克拉美-罗界为

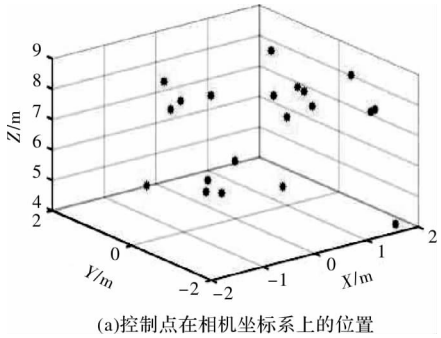
$$\boldsymbol{C}_{CR}(\theta)=\frac{\delta_w^2}{2K}\{\mathrm{Re}[\boldsymbol{D}\boldsymbol{D}^{\mathrm{H}}]\}^{-1} \tag{30}$$

从(30)式中可以看出，参数估计的克拉美-罗界随拍摄数量的增加线性下降，随噪声功率的增加线性增加。

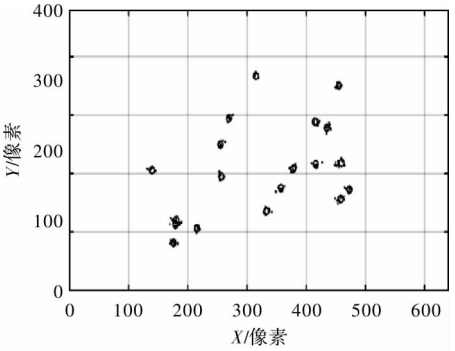
4 仿真

利用计算机仿真对最大似然估计和克拉美-罗界进行分析，并与传统迭代算法以及 OI 算法进行比较。

仿真参数：相机焦距 $f=35$ mm，单个像素尺寸 $d_x=d_y=62.5$ μm，主点在图像中心，控制点从相机坐标系的$[-2\quad 2]\times[-2\quad 2]\times[4\quad 9]$ m^{3}的长方形区域内随机选取，相机坐标系相对于世界坐标系的欧拉角为$[20\quad 10\quad 30]°$，平移矩阵为$[2\quad 3\quad 10]$ m。图 2(a)是随机选取的控制点在相机坐标系下的坐标，图 2(b)中圆圈表示相机拍摄控制点时实际成像位置(像素坐标系)，圆圈旁边的 10 个点为噪声功率为 1 dB 时，由拍摄的 10 张图片提取的坐标点。



(a)控制点在相机坐标系上的位置

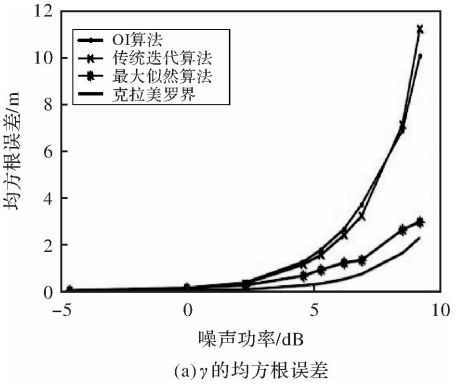


(b)控制点在像平面上的位置

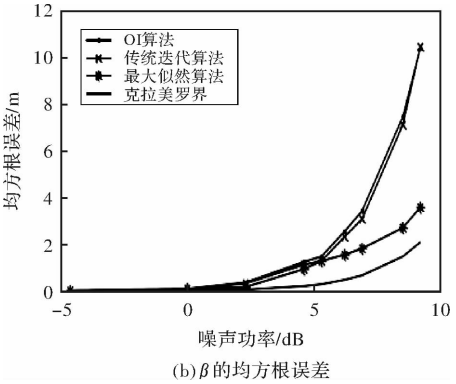
图 2 控制点在相机坐标系和成像平面上的位置

Fig. 2 Positions of control points in camera coordinate system and image plane

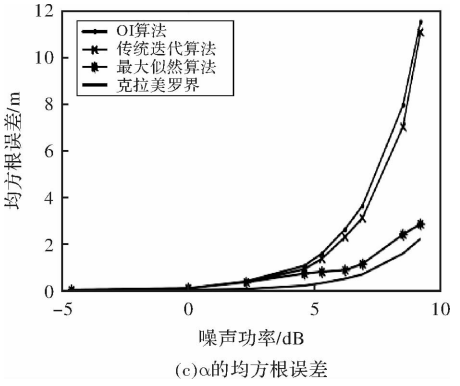
图 3 是传统迭代算法、OI 算法、最大似然方法在不同的噪声功率下的参数估计均方根误差(root-mean-square error, RMSE)，在各噪声级下完成 500 次独立实验，其中最大似然方法利用了 10 张图片。从仿真结果中可以看出，3 种方法在噪声功率较小的情况下都趋向于克拉美-罗界，传统迭代算法性能略优于 OI 算法，而最大似然方法的性能明显优于其他两种方法。



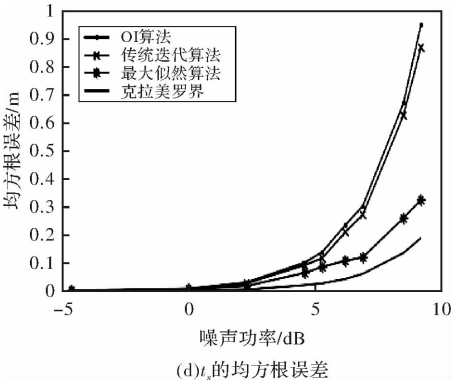
(a)γ的均方根误差



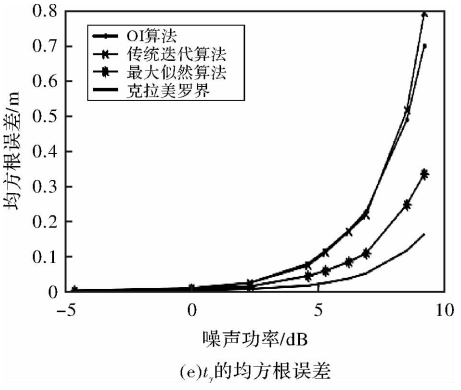
(b)β的均方根误差



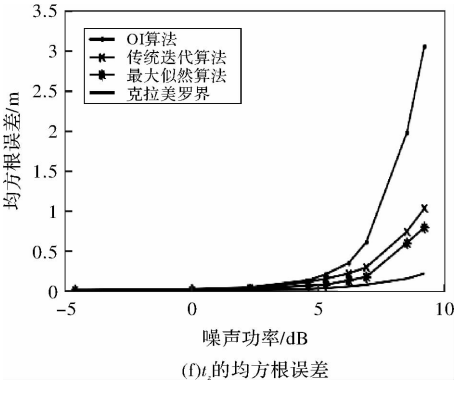
(c)α的均方根误差



(d)t_x的均方根误差



(e)t_y的均方根误差

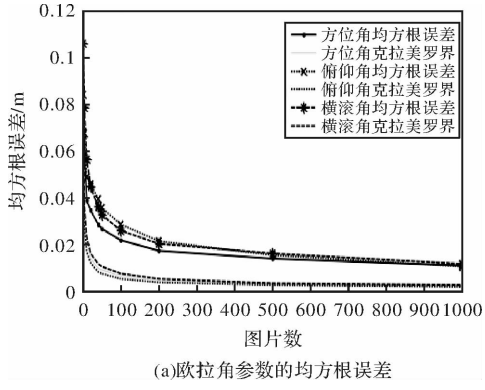


(f)t_z的均方根误差

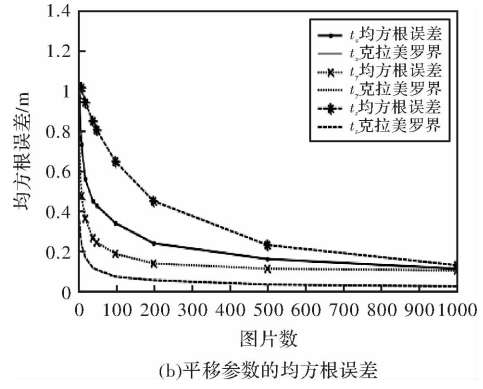
图 3　不同噪声功率下各算法的参数估计均方根误差

Fig. 3　RMSEs of different methods with different noise powers

图 4 是最大似然估计在不同的图片数下的参数估计均方根误差，噪声功率为 2 dB，每个图片数完成 500 次独立实验，可以看出参数估计的 RMSE 随着图片数的增加而减小，说明适当地增加拍摄数量可有效提高估计的性能。



(a)欧拉角参数的均方根误差



(b)平移参数的均方根误差

图 4　不同图片数下参数估计的均方根误差

Fig. 4　RMSEs with different snapshot numbers

5　结论

给出机器视觉位姿估计的信号模型，并从随机信号的角度出发，推导了基于机器视觉的最大似然位姿估计以及相应的克拉美-罗界，从理论上证明了利用单幅图像时，在各向同性高斯噪声情况下传统迭代算法与最大似然算法等效，通过仿真分析，可知在图像数量适当增加的情况下，估计性能得到了明显改进，适用于更高精度的位姿估计。并且推导的位姿估计克拉美-罗界作为任何无偏估计的方差下界，是位姿估计的性能极限，可以作为位姿估计方法有效性的评价标准；下一步，针对复杂的噪声情况，应在此基础上讨论各向异性噪声以及相关噪声情况下的位姿估计方法及其克拉美-罗界。

参考文献：

[1]　LI S Q, XU C. Efficient lookup table based camera

pose estimation for augmented reality[J]. Computer Animation and Virtual Worlds, 2011, 22(1): 47-58.

[2] CAMPA G, MAMMARELLA M, NAPOLITANO M R, et al. A comparison of pose estimation algorithms for machine vision based aerial refueling for UAV[C]. US: IEEE, 2006.

[3] WANG Qiyue, WANG Zhongyu. Position and pose measurement of spacecraft based on monocular vision [J]. Journal of Applied Optics, 2017, 38(2): 250-255.
汪启跃, 王中宇. 基于单目视觉的航天器位姿测量[J]. 应用光学, 2017, 38(2): 250-255.

[4] LI S Q, XU C, XIE M. A robust O(n) solution to the perspective-n-point problem[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1444-1450.

[5] LEPETIT V, MORENO-NOGUER F, FUA P. EPnP: an accurate O(n) solution to the PnP problem[J]. International Journal of Computer Vision, 2009, 81(2): 155-166.

[6] HMAM H, KIM J. Optimal non-iterative pose estimation via convex relaxation[J]. Image and Vision Computing, 2010, 28(11): 1515-1523.

[7] LOWE D G. Three-dimensional object recognition from single two-dimensional images[J]. Artificial Intelligence, 1987, 31(3): 355-395.

[8] CHEN Peng. Monocular based camera pose estimation[D]. Beijing: University of Science and Technology Beijing, 2015.
陈鹏. 基于单目视觉的像机位姿估计技术[D]. 北京: 北京科技大学, 2015.

[9] LU C P, HAGER G D, MJOLSNESS E. Fast and globally convergent pose estimation from video images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(6): 610-622.

[10] LI Xin, LONG Gucan, LIU Jinbo, et al. Accelerative orthogonal iteration algorithm for camera pose estimation[J]. Acta Optica Sinica, 2015, 35(1): 258-265.
李鑫, 龙古灿, 刘进博, 等. 相机位姿估计的加速正交迭代算法[J]. 光学学报, 2015, 35(1): 258-265.

[11] SCHWEIGHOFER G, PINZ A. Robust pose estimation from a planar target[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(12): 2024-2030.

[12] LUO Pengfei, ZHANG Wenming, LIU Zhong, et al. Fundamentals of statistical signal processing[M]. Beijing: Publishing House of Electronics Industry, 2014: 155-202.
罗鹏飞, 张文明, 刘忠, 等. 统计信号处理基础[M]. 北京: 电子工业出版社, 2014: 155-202.

[13] STOICA P, NEHORIA A. Mode, maximum likelihood, and Cramér-Rao bound: conditional and unconditional results[R]. New Haven: Center for Systems Science Yale University, 1989.

[14] FRIEDLANDER B, PORAT B. The exact Cramer-Rao bound for Gaussian autoregressive processes[J]. IEEE Transactions on Aerospace and Electronic Systems, 1989, 25(1): 3-7.

[15] VAN TREES H L. Detection, estimation, and modulation theory, Part VI[M]. New York: Wiley Interscience, 2001: 689-785.