

一种三维刚体运动参数鲁棒估计的神经算法¹

陈海峰 邵家玉 陈维南
(东南大学自动化研究所 南京 210096)

摘 要 该文提出一种基于特征点匹配的刚体运动参数估计方法。在运动估计线性算法的基础上,文中利用全最小二乘(TLS)方法来进行求解,并建立次分量提取神经元来获得该全最小二乘解。基于测量数据中出格点(Outlier)的存在,我们在神经元的权值学习规则中引入鲁棒估计思想。实验结果表明,该方法能有效地克服出格点产生的误差,准确地估计出刚体的三维运动参数,比较令人满意。

关键词 刚体, 运动, 参数估计, 神经元, 鲁棒, 全最小二乘法, 次分量

中图分类号 TNP391

1 引 言

三维刚体运动参数估计是运动视觉中的一个重要领域,其问题可表达为:给定相继时刻 t_1, t_2 的两幅图像 $P_1(x, y), P_2(x, y)$,求图中刚体的运动参数 (R, t) ,其中 R 是刚体的旋转矩阵, t 是平移向量。该问题的研究方法大致可分为基于特征点匹配的方法和基于光流场的方法两类。本文所研究的是基于特征点匹配的参数估计方法,它包括两个步骤,首先从相继两幅图像中抽取特征点并建立起对应,接着根据这些特征点之间的对应来计算物体的运动参数。这里我们讨论如何根据已知的特征点求运动参数,对于特征点间的抽取及对应问题可参看文献[1,2]。

文献中已有许多关于根据特征点对应求运动参数的成熟算法。其中著名的有 Tai & Huang 的线性算法^[3],但是该方法的最大缺点就是对噪声非常敏感。另外在实际问题中,刚体运动前后两幅图像的特征点间可能发生匹配错误,甚至混入一些非刚体上的点,这些点称为“出格点”(Outliers)。为此人们提出了各种各样的鲁棒算法,如 Haralick 提出的 M 估计方法^[4],庄新华等提出的模型匹配估计算法^[5],谢嘉淋、吴立德提出的最小平方中值估计算法^[6]等,这些方法都不同程度地克服了出格点带来的误差。本文中我们提出一种刚体运动参数鲁棒估计的神经元模型,首先我们将刚体的运动参数估计转化为求解一线性超定方程组,并指出该方程组的全最小二乘解要优于最小二乘解,其值等于特征点数据协方差矩阵最小特征值对应的特征向量,于是利用次分量提取神经元(MCA)来获得该全最小二乘解。考虑到特征点匹配中错误匹配的存在,我们将鲁棒估计的思想引入神经元的学习规则中。实验结果表明我们的方法能够有效克服噪声及出格点的影响,准确地估计出刚体的运动参数。

2 刚体的运动参数估计

我们对摄像机采用透视投影模型,设焦距 $f = 1$,那么摄像机中心坐标系 (X, Y, Z) 和图像坐标系 (x, y) 之间的关系为 $x = X/Z, y = Y/Z$ 。在摄像机中心坐标系中,刚体的运动关系可以用 3×3 的旋转矩阵 R 和平移向量 t 来表示:

$$[X'Y'Z']^T = R \cdot [XYZ]^T + t, \quad (1)$$

其中

¹ 1998-10-08 收到, 1999-05-12 定稿

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}.$$

这里我们已经在连续两帧图像 $P_1(x, y)$, $P_2(x, y)$ 上提取了在同一刚体上的点 $(x_1, y_1), (x_2, y_2) \cdots (x_n, y_n)$ 和 $(x'_1, y'_1), (x'_2, y'_2) \cdots (x'_n, y'_n)$, 并建立起它们之间的对应 $(x_i, y_i) \leftrightarrow (x'_i, y'_i)$, 问题在于如何根据上述 n 组对应求刚体的运动参数 (R, t) .

由 n 组对应点求运动的方法很多, 其中较有影响的是 Tai&Huang 的线性算法^[3], 该方法引入一反对称矩阵 $T = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}$, 并令 $E = T \cdot R = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}$, 显然有 $T \cdot t = 0$, 并对任意向量 u , 有 $u^T T u = 0$. 在 (1) 式两边同时乘以 $(X'Y'Z')^T T$, 则最终整理得

$$(x, y, 1) \cdot E \cdot (x', y', 1)^T = 0. \quad (2)$$

线性方程 (2) 式是齐次的, 因而独立参数只有 8 个, 这里不妨设 $h_9 = -1$, 那么 (2) 式可以表示为

$$a^T h = 1, \quad (3)$$

其中 $a = (xx', yy', x'y', yx', y'y', y'x', x, y)^T$, $h = (h_1, h_2, h_3, \dots, h_8)^T$. 每一组特征点对 $(x_i, y_i) \leftrightarrow (x'_i, y'_i)$, 对应于一条线性方程 $a_i^T h = 1$, 可见要求得未知参数 h , 至少要有 8 组对应点. 但由于噪声的存在, 一般都要多于 8 组对应点来求解, 这就构成一个超定方程组:

$$Ah = e, \quad (4)$$

其中 $A = (a_1, a_2, \dots, a_n)^T$, $e = (1, 1, \dots, 1)^T$. 求得参数 h 后, 就很容易求得刚体的运动参数 (R, t) ^[3].

对于方程组 (4) 式, 许多文献都是采用最小二乘法 (LS) 求解^[3,7], 即求使得 $\|e - Ah\| = \min$ 的向量 h , 但最小二乘法只是通过 e 在 $R(A)$ 上的最佳逼近 e_{LS} 来获得解, 它仅考虑了观测向量 e 的不精确性, 但实际问题中数据矩阵 A 也存在误差. 为此我们采用全最小二乘法 (TLS) 来估计 h , 即使

$$\|e - Ah\|/\|h\| = \min. \quad (5)$$

文献 [8] 中详细论述了全小二乘法的原理及性质, 并指出全最小二乘法对数据的拟合残差要小于最小二乘解的残差, 我们的实验也证明了这一点.

由 (5) 式得, TLS 拟合问题就是使下式的全最小平方误差 E 最小:

$$E = \frac{1}{n} \sum_{i=1}^n r_i^2 = \frac{1}{n} \sum_{i=1}^n \frac{(h^T a_i - 1)^2}{h^T h} = \frac{h^T R h - 2h^T u + 1}{h^T h}, \quad (6)$$

其中 $R = \frac{1}{n} \sum_{i=1}^N a_i a_i^T$, $u = \frac{1}{n} \sum_{i=1}^N a_i$. 另外对 (3) 式两边取期望, 有 $h^T u = 1$, 那么有

$$E = \frac{h^T R h - h^T u u^T h}{h^T h} = \frac{h^T \phi h}{h^T h}, \quad (7)$$

其中 $\phi = R - uu^T$.

可见 (7) 式中的 TLS 解 h 等于数据协方差矩阵 ϕ 的最小特征值所对应的归一化特征向量。这里我们建立次分量分析神经元来抽取该特征向量, 如图 1 所示。

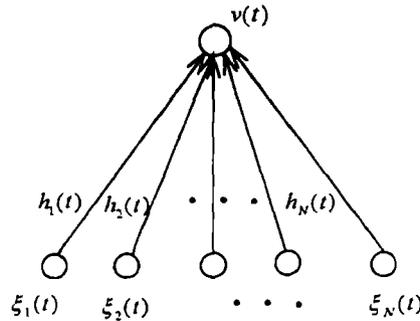


图 1 归一化反 Hebb 神经元

神经元的输出为 $v(t)$, 输入为 $\xi(t) = a - u = (\xi_1(t), \xi_2(t), \dots, \xi_N(t))^T$, 权值 $h(t) = (h_1(t), h_2(t), \dots, h_N(t))^T$, $N=8$. 利用归一化反 Hebb 规则对神经元权值进行训练,

$$\begin{aligned} h(t+1) &= h(t) - \eta(t) \cdot v(t) \cdot \left[\xi(t) - \frac{v(t)h(t)}{h^T(t)h(t)} \right] \\ &= h(t) - \eta(t) \cdot \left[\xi(t)\xi^T(t)h(t) - \frac{h^T(t)\xi(t)\xi^T(t)h(t)}{h^T(t)h(t)} h(t) \right], \end{aligned} \quad (8)$$

其中 $\eta(t)$ 是学习速度因子。

可以用随机逼近论的方法来分析上述学习规则。不难说明, 如果假设 $\xi(t)$ 的分布满足某些较弱的条件^[9]且学习增益 $\eta(t)$ 以适当方式减少到零 (例如正比于 $1/t$), 那么可用下面微分方程来近似 (8) 式

$$\frac{dh(t)}{dt} = -[\xi(t)\xi^T(t)h(t) - \frac{h^T(t)\xi(t)\xi^T(t)h(t)}{h^T(t)h(t)} h(t)]. \quad (9)$$

如果进一步假设输入矢量 $\xi(t)$ 在整个学习阶段保持平稳, 且 $\xi(t)$ 与 $h(t)$ 统计不相关, 那么对 (9) 式两边取平均得相应的平均化微分方程:

$$\frac{dh(t)}{dt} = -\phi h(t) + \frac{h^T(t)\phi h(t)}{h^T(t)h(t)} h(t). \quad (10)$$

(10) 式的渐近轨迹和 (9) 式的渐近轨迹以很大的概率相接近, 最终 (9) 式的解以概率 1 分别趋于 (10) 式的均匀渐近稳定解。

定理 1 设 ϕ 是一个具有重数为 1 的最小特征值的正半定矩阵, 并令 λ_{\min} 和 C_{\min} 分别是 ϕ 的最小特征值及相应的归一化特征矢量, 如果 $h^T C_{\min}(0) \neq 0$, 且 (10) 式成立, 则

$$\lim_{t \rightarrow \infty} h(t) = C_{\min}, (\text{或 } -C_{\min}), \quad \lim_{t \rightarrow \infty} h(t)\phi h(t) = \lambda_{\min} = \min_h \{h^T \phi h\}.$$

该定理的证明见文献 [9], 图 1 中的神经元根据规则 (8) 式学习后, 其权矢量就对应于满足 (5) 式最小的矢量 h , 并由此获得运动参数 (R, t) 。

3 出格点的影响及解决措施

由上面的叙述可知, 运动前后图像的每一组特征点对对应着一数据矢量 a_i , 而 n 个数据矢量 $\{a_i, i = 1, 2, \dots, n\}$ 的协方差矩阵 ϕ 最小特征值对应的特征矢量 h 包含着运动信息。根据次分量提取规则 (8) 式求得 h , 即可获得运动参数 (R, t) 。

但在实际问题中, 运动前后图像的特征点对一般都是有误差的。这些误差大致分为两类, 一类是特征点本身定位的误差, 它们一般比较小, 可以用正态噪声来表示。另一类则是由于出现假的特征点或是特征点之间产生错误的匹配而造成的, 它们一般比较大, 不能用正态噪声的模型来表示, 在统计中称为“出格点”。出格点对 (8) 式是计算结果的影响是比较大的, 因此很有必要引入鲁棒估计的思想, 对 (8) 式中的学习规则加以改进, 以减少“出格点”对解的影响。

考虑 (7) 式中的能量函数:

$$E = \frac{h^T \phi h}{h^T h} = \frac{1}{n} \sum_{i=1}^n \frac{h^T \xi_i \xi_i^T h}{h^T h}. \quad (11)$$

其中 $\xi_i = a_i - u$, $i = 1, 2, \dots, n$ 。由 (11) 式, 我们认为每一组特征点对 i 对能量函数 E 的贡献大小为 $Z(\xi_i, h) = \frac{h^T \xi_i \xi_i^T h}{h^T h}$, 则

$$E = \frac{1}{n} \sum_{i=1}^n Z(\xi_i, h). \quad (12)$$

超定方程组 (4) 式可视为八维空间的超平面拟合, 每组数据 a_i 对应于该空间的一个采样点, 最小化能量函数 E 获得的超平面能够有效地逼近于每个采样点。但是由于采样点中出格点的存在, 我们不再要求拟合的超平面去逼近每一个采样点。而是要求该超平面有足够的“力量”逼近于系统的非出格采样点。为此, 引入一组二值变量 $V = \{V_i, i = 1, 2, \dots, n\}$, $V_i = 1$ 表示第 i 组特征点对的数据矢量为非出格数据, 否则 $V_i = 0$ 。于是定义新的能量函数:

$$E = \sum_{i=1}^n V_i \cdot Z(\xi_i, h) + \alpha \sum_{i=1}^n (1 - V_i). \quad (13)$$

(13) 式右边后半部分表示我们对 $\{V_i\}$ 分布的先验假设, 与 (12) 式的能量函数相比, (13) 式能够根据 $Z(\xi_i, h)$ 的大小来判断数据点 i 是否为出格点。当 $Z(\xi_i, h) > \sqrt{\alpha}$ 时, 则将 ξ_i 不予考虑到能量函数 E 中。

直接求 (13) 式的极小是很困难的, 我们将求 (13) 式的极小转化为求下式的极大:

$$P(V, h) = (1/Z) \exp[-\beta J(V, h)]. \quad (14)$$

(14) 式认为变量 V , h 的分布为 Gibbs 分布, 其中 Z 是“配分函数” (Partition Function): $Z = \sum_V \int_h P(V, h) dh$ 。下面求 $P(V, h)$ 关于 h 的边缘分布 $P_{\text{margin}}(h)$ 。

$$\begin{aligned}
P_{\text{margin}}(h) &= (1/Z) \sum_v \exp\{-\beta \sum_i (V_i \cdot Z(\xi_i, h) + \alpha(1 - V_i))\} \\
&= (1/Z) \prod_i \sum_{V_i \in \{0,1\}} \exp\{-\beta \sum_i (V_i \cdot Z(\xi_i, h) + \alpha(1 - V_i))\} \\
&= (1/Z) \prod_i \{\exp(-\beta \cdot Z(\xi_i, h)) + \exp(-\beta\alpha)\} \\
&= (1/Z) \exp(-n\beta\alpha) \prod_i \{1 + \exp(-\beta(Z(\xi_i, h) - \alpha))\}.
\end{aligned}$$

定义 $Z_h = Z \cdot \exp(n\beta\alpha)$, 则得到 $P_{\text{margin}}(h) = (1/Z_h) \exp(-\beta E_{\text{eff}}(h))$, 其中

$$E_{\text{eff}}(h) = (1/\beta) \sum_i \{1 + \exp(-\beta(Z(\xi_i, h) - \alpha))\}. \quad (15)$$

我们希望求得 h' , 它使得 $\{V, h\}$ 的联合分布 $P(V, h)$ 最大。但考虑到问题的复杂性, 这里用使边缘分布 $P_{\text{margin}}(h)$ 最大的 h 来作为 h' 的近似。而 $P_{\text{margin}}(h)$ 的求极大等效于极小化能量函数 $E_{\text{eff}}(h)$, 由于函数 $E_{\text{eff}}(h)$ 的非凸性, 我们采用统计物理中的“确定退火”(deterministic annealing) 的方法来极小化 $E_{\text{eff}}(h)$ 。首先取较小的 β 值, 利用梯度下降规则 (16) 式来极小化 $E_{\text{eff}}(h)$, 接着不断增大 β , 直至无穷大, 而对应的 $E_{\text{eff}}(h)$ 也最终趋于全局最小。

利用梯度下降法求 $E_{\text{eff}}(h)$ 的极小, 有 $\partial h / \partial t = -\partial E_{\text{eff}}(h) / \partial h$, 从而有

$$\begin{aligned}
h(t+1) &= h(t) - \eta'(t) \partial E_{\text{eff}}(h(t)) / \partial h(t) \\
&= h(t) - \eta'(t) \sum_i \frac{1}{1 + \exp\{-\beta(Z(\xi_i, h(t)) - \alpha)\}} \frac{\partial Z(\xi_i, h(t))}{\partial h(t)} \\
&= h(t) - \eta'(t) \sum_i \frac{1}{1 + \exp\{-\beta(Z(\xi_i, h(t)) - \alpha)\}} \\
&\quad \times \left[\xi_i \xi_i^T h(t) - \frac{h_{(t)}^T \xi_i \xi_i^T h(t)}{h_{(t)}^T h(t)} h(t) \right]. \quad (16)
\end{aligned}$$

规则 (16) 式是特征点集的平均化迭代形式。我们也可以用每一组特征点对应的能量函数进行迭代求解, 即将 (16) 式改写与 (8) 式相类似的形式:

$$\begin{aligned}
h(t+1) &= h(t) - \eta'(t) \frac{1}{1 + \exp\{-\beta(Z(\xi_i, h(t)) - \alpha)\}} \\
&\quad \times \left[\xi_{(t)} \xi_{(t)}^T h(t) - \frac{h_{(t)}^T \xi_{(t)} \xi_{(t)}^T h(t)}{h_{(t)}^T h(t)} h(t) \right]. \quad (17)
\end{aligned}$$

其中 $Z(\xi(t), h(t)) = \frac{h_{(t)}^T \xi_{(t)} \xi_{(t)}^T h(t)}{h_{(t)}^T h(t)}$ 。

与 (8) 式相比, 算法 (17) 式增加了一项 $\frac{1}{1 + \exp\{-\beta(Z(\xi_i, h(t)) - \alpha)\}}$ 。我们认为 (17) 式是规则 (8) 式中引入鲁棒 M 估计算子^[10]后产生的, 实验结果表明其有效地抑制了出格数据对解的影响。

4 实验与仿真结果

这里我们进行了实验,将文献[3]中的线性算法,(8)式对应的次分量提取算法(MCA),及(17)式对应的鲁棒次分量提取方法(RMCA)进行了比较。

我们用围绕 x, y, z 三个坐标轴的旋转角 φ, θ, ψ 来表示旋转矩阵 R , 这里取 $\varphi = 15^\circ$, $\theta = 10^\circ$, $\psi = 20^\circ$, 平移向量 $t = (2, 2, 2)$ 。刚体运动前的 3D 点在 $(-6, -6, 3)$ 和 $(6, 6, 6)$ 内随机产生,用(1)式可计算出其运动后的三维坐标,对于运动前后的三维坐标进行透视投影,便得到两维投影坐标。

对运动后的两维投影坐标加均值为 0 的 Gauss 噪声,信噪比 $\text{SNR}=10$, 另外,用随机生成的投影坐标来取代部分运动后的投影坐标,使其成为“出格点”。这里我们取特征点数 $n=50$, 出格点所占的比例 $\rho = 10\%$ 。

实验中有三个参数需要确定。首先是(8)式和(17)式的学习增益 $\eta(t)$, 该参数影响算法的收敛性及收敛速度。在实验中,我们将 $\eta(t)$ 从初始值 $\eta(0) = 1.0$ 出发,随着迭代的进行不断减小,其减小速度与 $1/t$ 成正比。另外需要确定的参数是(17)式中的退火因子 β 和阈值参数 α 。为了避免算法收敛于能量函数的局部极值,退火因子 β 在迭代的过程中是可变的,其初始一般取得足够小,随着迭代的进行 β 不断增加,增加速度与 $\ln t$ 成正比。阈值参数 α 的选取则是根据我们对结果的经验期望,取 $\alpha=1.4$ 。

实验结果如图 2 所示,图 2(a)-2(d) 中的横坐标表示算法的迭代次数,图 2(a), 2(b), 2(c) 分别表示所估计的旋转角 φ, θ, ψ 的大小,图 2(d) 表示所估计的平移向量与真实平移向量之间的夹角 $\Delta\alpha$ 。图中与 LS 相对应的虚线为线性算法的结果, MCA 表示的迭代轨迹为(8)式对应的次分量提取算法的结果, RMCA 表示的迭代轨迹为(17)式对应的鲁棒次分量提取算法的结果。

图 2 的结果告诉我们,次分量提取算法 MCA 要比线性算法更有效地抑制噪声,但它无法解决“出格点”对解的影响。相对而言,鲁棒次分量提取算法 RMCA 不仅能够有效地抑制 Gauss 噪声,而且能有效地除去出格点,从而得到运动参数稳定的估计值,是令人满意的。

5 结 论

在运动估计线性算法的基础上,本文在三维刚体运动参数估计方面作了以下几点创新:

(1) 我们采用全最小二乘法,而不是一般的最小二乘法来求解线性方程组,并指出全最小二乘法获得的解要优于最小二乘解,从而得到更加准确的运动参数估计。

(2) 考虑到神经网络较强的自学习及并行处理能力,文中建立次分量提取神经元(MCA)来获得方程组的全最小二乘解,进一步提高了算法的效率。

(3) 由于二维特征点中错误匹配的存在,我们将鲁棒估计的思想引入神经元的学习规则中,从而抑制了错误匹配对最终结果的影响。

实验结果表明,本文提出的三维刚体运动参数鲁棒估计的神经算法要比原有的线性算法结果更加准确,鲁棒性更强,并且计算效率高,是比较令人满意的。

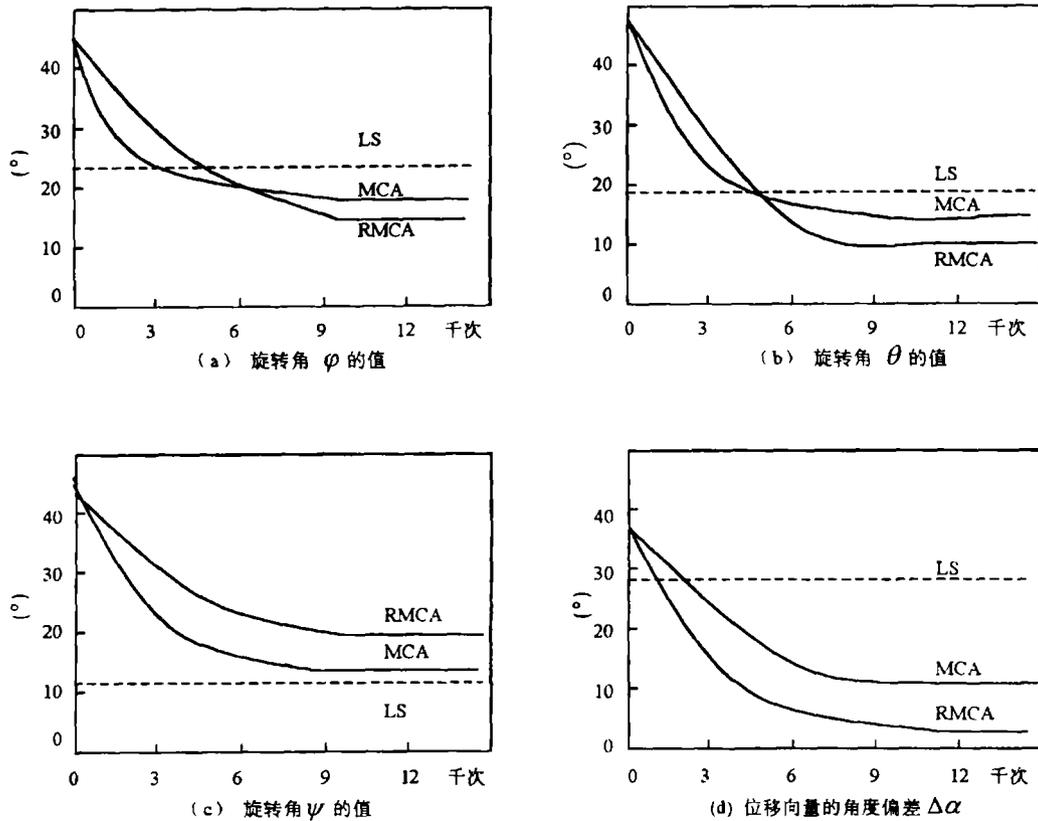


图2 实验结果

参 考 文 献

- [1] Lee C H. Correspondence problem in image sequence analysis. *Pattern Recognition*, 1993, 26(1): 47-61.
- [2] Sethi I K, Patel N V, *et al.* A general approach for token correspondence. *Pattern Recognition*, 1994, 27(12): 1775-1786.
- [3] Tai R Y, Huang T S. Uniqueness and estimation of 3D motion parameters of rigid bodies with curved surfaces, *IEEE Trans. on PAMI*, 1984, PAMI-6(1): 13-27.
- [4] Haralick R M, Lee J H, *et al.* Pose estimation from correspondence point data. *IEEE Trans. on SMC*, 1989, SMC-19(6): 1426-1446.
- [5] 庄新华等. 刚体空间运动参数的鲁棒估计. *计算机学报*, 1991, 14(9): 660-670.
- [6] 朱嘉琳, 吴立德. 用 2D 点匹配恢复刚体运动参数的鲁棒估计. *计算机学报*, 1994, 17(5): 388-392.
- [7] Wang J, Huang T S, *et al.* Motion and structure from two perspective views: algorithms, error analysis, and error estimation. *IEEE Trans. on PAMI*, 1989, PAMI-11(5): 451-475.
- [8] 张洪斌等. 全最小二乘法及其在参数估计中的应用. *自动化学报*, 1995, 21(1): 40-47.
- [9] Oja E, Karhunen J. On stochastic approximation of eigenvectors and eigenvalues for the expectation of a random matrix. *J. Math. Anal. Appl.* 1985, 16(1): 69-84.

- [10] Serfling R. J. Approximation Thereoms of Mathematical Statistics. New York: John Wiley & Sons Inc, 1980, 243-261.

A NEURAL ALGORITHM FOR ESTIMATING THE 3D RIGID BODIES' MOTION PARAMETERS

Chen Haifeng Shao Jiayu Chen Weinan

(Institute of Automation, Southeast University, Nanjing 210096)

Abstract This paper deals with estimating the rigid bodies' motion parameters from token correspondence between two perspective views. Total least square(TLS) method is used to solve the linear equations of motion estimation. Then a minor component analysis(MCA) neuron is constructed to get the TLS solution. Finally a outline resisting version of the learning rule is developed by using the statistical approach. Comparative experiments have been made and results show that our robust rules improve the performance significantly when outliers are presented.

Key words Rigid body, Motion, Parameter estimation, Neuron, Robust, Total least square, Minor component

陈海峰: 男, 1975 年生, 博士, 研究方向为图像处理和机器视觉.

邵家玉: 男, 1971 年生, 博士, 主要从事数据库与人工智能的研究.

陈维南: 男, 1936 年生, 教授, 博士生导师, 江苏省自动化学会理事长, 长期从事智能控制和机器视觉的研究.